CANADIAN
JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. XIII - NO. 3
1961

Math
THE UNIVERSITY OF MICHIGAN
JUL 12 1961
MATHEMATICS LIBRARY

Matrix commutators M. F. Smiley 353

Orthomorphisms of groups and orthogonal latin squares. I
Diane M. Johnson, A. L. Dulmage, and N. S. Mendelsohn 356

Partition rings of cyclic groups of odd prime power order
K. I. Appel 373

On the structure of semi-prime rings and their rings of quotients
Joachim Lambek 392

Some two-dimensional unitary groups generated by three reflections
D. W. Crowe 418

Moulton planes William A. Pierce 427

Tetraspheres. I A. de Majo 437

Polar means of convex bodies and a dual to the Brunn-Minkowski theorem William J. Firey 444

On the Hausdorff and trigonometric moment problems P. G. Rooney 454

The expansion problem with boundary conditions at a finite set of points Randal H. Cole 462

On stability in the large for systems of ordinary differential equations Philip Hartman 480

Asymptotic solutions of equations in Banach space
C. A. Swanson and M. Schulzer 493

Normal operators on the Banach space $L^p(-\infty, \infty)$. Part I Gregers L. Krabbe 505

Homogeneous continua which are almost chainable
C. E. Burgess 519

Published for

THE CANADIAN MATHEMATICAL CONGRESS

by the

University of Toronto Press

EDITORIAL BOARD

H. S. M. Coxeter, G. F. D. Duff, R. D. James, R. L. Jeffery, J.-M. Maranda, G. de B. Robinson, P. Scherk

with the co-operation of

D. B. DeLury, J. Dixmier, W. Fenchel, H. Freudenthal, I. Kaplansky, N. S. Mendelsohn, C. A. Rogers, H. Schwerdtfeger, A. W. Tucker, W. J. Webber, M. Wyman

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, G. F. D. Duff, University of Toronto. Authors are asked to write with a sense of perspective and as clearly as possible, especially in the introduction. Regarding typographical conventions, attention is drawn to the Author's Manual of which a copy will be furnished on request.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is $10.00. This is reduced to $5.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta	Assumption University
University of British Columbia	Carleton University
Dalhousie University	Ecole Polytechnique
Université Laval	Loyola College
University of Manitoba	McGill University
McMaster University	Université de Montréal
Mount Allison University	Nova Scotia Technical College
Queen's University	St. Mary's University
University of Saskatchewan	University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

MATRIX COMMUTATORS

M. F. SMILEY

Introduction. A classical theorem states that if a square matrix B over an algebraically closed field F commutes with all matrices X over F which commute with a matrix A over F, then B must be a polynomial in A with coefficients in F **(2)**. Recently Marcus and Khan **(1)** generalized this theorem to double commutators. Our purpose is to complete the generalization to commutators of any order.

Let F be an algebraically closed field and let F_n be the ring of all n by n matrices with elements in F. We define $\Delta_Y Z = [Z, Y] = ZY - YZ$ for all Y, Z in F_n.

THEOREM. *Let $A, B \in F_n$ be such that for some positive integer s, $\Delta_A{}^s X = 0$ for X in F_n implies that $\Delta_X{}^s B = 0$. Let the characteristic of F be* 0 *or at least n. Then B is a polynomial in A with coefficients in F.*

For $s = 1$ we have the classical theorem except for the restriction on the characteristic of F. For $s = 2$ we have the result of Marcus and Khan with a bit more freedom for the characteristic of F. We feel that even for $s = 2$ our proof has interest. We first observe that $s > 1$ is "rather without meaning" for semi-simple matrices and then we use this observation to reduce our theorem to the classical case. Here we call A in F_n semi-simple in case the roots of the minimal polynominal of A are distinct.

1. Some lemmas. The results of this section will be used in the next section in which we will prove our theorem.

LEMMA 1. *If A is semi-simple in F_n, then $\Delta_A{}^s X = 0$ for some positive integer s only if $\Delta_A X = 0$.*

Proof. We use induction on s. Let $E_k (k = 1, \ldots, q)$ be the principal idempotents of A so that $A = \mu_1 E_1 + \ldots + \mu_q E_q$ with $\mu_k \in F$ $(k = 1, \ldots, q)$. Then each E_k is a polynomial in A with coefficients in F. The Jacobi identity $[Y, [Z, W]] + [Z, [W, Y]] + [W, [Y, Z]] = 0$ for all Y, Z, W in F_n shows that if $E = E_k$ $(k = 1, \ldots, q)$, then $\Delta_A \Delta_E Y - \Delta_E \Delta_A Y = 0$ for all Y in F_n. Now $\Delta_A{}^s X = [\Delta_A{}^{s-1} X, A] = 0$ gives $[\Delta_A{}^{s-1} X, E] = 0$ and hence $\Delta_A{}^{s-1} \Delta_E X = 0$. By our inductive hypothesis, $\Delta_A \Delta_E X = 0$ from which $\Delta_E{}^2 X = 0$ follows at once. But $\Delta_E{}^2 X = 2EXE + XE - EX = 0$ yields $EX = XE$ upon right and left multiplication by E. Thus $\Delta_E X = 0$ for all $E = E_k$ $(k = 1, \ldots, q)$ and consequently $\Delta_A X = 0$, completing our inductive proof of the lemma.

Received May 12th, 1960.

An alternative proof of Lemma 1 is suggested by the referee. We may assume that A is a diagonal matrix and use the well-known matrix representation $L = I \otimes A - A \otimes I$ for Δ_A, where $\otimes$ denotes the Kronecker product. But then L is a diagonal matrix so that L and L^s have the same null-space, and this proves Lemma 1.

At this point we introduce the usual matrix units e_{ij} $(i, j = 1, \ldots, k)$ in F_k. The matrix e_{ij} has 1 in the ith row and jth column and zeros elsewhere.

LEMMA 2. *In F_k, let $C = \lambda I_k + e_{21} + e_{32} + \ldots + e_{kk-1}$ with λ in F and $X = e_{11} + 2e_{22} + \ldots + ke_{kk}$. Then $\Delta_C{}^2X = 0$, and for $Y = (C - \lambda)X$, $\Delta_C{}^2Y = 0$.*

Proof. A simple computation shows that $\Delta_C X = XC - CX = C - \lambda I_k$. Since $\Delta_C(C - \lambda)T = (C - \lambda)\Delta_C T$ for all T in F_k, the lemma follows. (The matrices X and Y are special cases of certain matrices used in (1) on pp. 273-274.)

LEMMA 3. *Let C, X, Y be as in Lemma 2 and let $B \in F_k$. Assume that the characteristic of F is 0 or at least k. Then $[B, X] = 0$ implies that B is a diagonal matrix and $[B, X] = [B, Y] = 0$ implies that B is a scalar matrix.*

Proof. With $B = \Sigma b_{ij}e_{ij}$ we find that $BX = \Sigma jb_{ij}e_{ij}$ and $XB = \Sigma ib_{ij}e_{ij}$. Hence $[B, X] = 0$ gives $b_{ij} = 0$ for $i \neq j$ and $i, j = 1, \ldots, k$. With $B = \text{diag}(b_1, \ldots, b_k)$, $YB = b_1e_{21} + 2b_2e_{32} + \ldots + (k - 1)b_{k-1}e_{kk-1}$ and $BY = b_2e_{21} + 2b_3e_{32} + \ldots + (k - 1)b_ke_{kk-1}$. Hence $[B, Y] = 0$ yields $b_1 = b_2 = \ldots = b_k$ so that B is a scalar matrix.

2. Proof of the theorem. In this section we use the lemmas of § 1 to prove our theorem. Since we shall use the classical result $(s = 1)$ in our proof, we assume that s is at least 2.

We may clearly assume that $A \in F_n$ is in Jordan normal form:

$$A = \text{diag}(C_1, \ldots, C_t) = \text{diag}(J_1, \ldots, J_q)$$

where each C_i $(i = 1, \ldots, t)$ is an n_i by n_i matrix corresponding to an elementary divisor $(x - \lambda_i)^{p_i}$ of A and each J_k is an m_k by m_k matrix with a single characteristic root μ_k and $\mu_k \neq \mu_l$ for $k \neq l$ $(k, l = 1, \ldots, q)$.

Take $X = \text{diag}(1, \ldots, n)$ and use Lemma 2 to obtain $\Delta_A{}^2X = 0$ and hence $\Delta_X{}^sB = 0$. By Lemma 1, since X is semi-simple, $\Delta_X B = 0$ and B must be diagonal by Lemma 3. We write $B = \text{diag}(B_1, \ldots, B_t)$, $X = \text{diag}(X_1, \ldots, X_t)$ conformally with $A = \text{diag}(C_1, \ldots, C_t)$. With $Y = \text{diag}((C_1 - \lambda_1)X_1, \ldots, (C_t - \lambda_t)X_t)$, we have $\Delta_A{}^2Y = 0$ by Lemma 2 and also $\Delta_A{}^2(X + Y) = 0$. Since $X + Y$ is semi-simple, $\Delta_{X+Y}B = \Delta_Y B = 0$. By Lemma 3, $B_i = c_iI_{n_i}$ with c_i in F $(i = 1, \ldots, t)$. Now let C_i and C_{i+1} have the same characteristic root λ and let U be an $(n_i + n_{i+1})$-rowed square matrix whose only non-zero element is 1 in the last row and first column. If $Z = \text{diag}(0, U, 0)$ in conformity with $A = \text{diag}(C_1, \ldots, C_t)$, then $ZA = AZ = \lambda Z$ so that

$\Delta_A Z = 0$. Since $X + Z$ is semi-simple, we obtain $\Delta_{X+Z} B = \Delta_Z B = 0$ from which $c_i = c_{i+1}$ follows. Thus if $B = \mathrm{diag}(B_{01}, \dots, B_{0q})$ in conformity with $A = \mathrm{diag}(J_1, \dots, J_q)$, then $B_{0k} = d_k I_{m_k}$ with d_k in F $(k = 1, \dots, q)$. Now if $[W, A] = 0$ it is well known that $W = \mathrm{diag}(W_1, \dots, W_q)$ in conformity with $A = \mathrm{diag}(J_1, \dots, J_q)$. A direct proof of this statement goes as follows. Partition W into blocks W_{kl} in conformity with $A = \mathrm{diag}(J_1, \dots, J_q)$. If $Y = W_{kl}$ with $k \neq l$, then $[W, A] = 0$ gives $(\rho I + C)Y = YD$ with C and D nil-potent and ρ non-zero in F. Thus $Y(R_D - R_C) = \rho Y$ where R_D, L_C denote right and left multiplications by C, D, respectively. Since C and D are nil-potent, so is $R_D - L_C$, and it follows that $\rho^t Y = 0$, $Y = 0$. Now we see that $[W, A] = 0$ for W in F_n implies that $[W, B] = 0$ and we complete the proof of our theorem by an appeal to the classical case.

References

1. M. Marcus and N. A. Khan, *On matrix commutators*, Can. J. Math., *12* (1960), 269–277.
2. J. H. M. Wedderburn, *Lectures on matrices*, Amer. Math. Soc. Colloq. Pub., *17* (New York, 1934).

University of Iowa
and
University of California, Riverside

ORTHOMORPHISMS OF GROUPS AND ORTHOGONAL LATIN SQUARES. I

DIANE M. JOHNSON, A. L. DULMAGE, AND N. S. MENDELSOHN

1. Introduction. Euler **(6)** in 1782 first studied orthogonal latin squares. He showed the existence of a pair of orthogonal latin squares for all odd n and conjectured their non-existence for $n = 2(2k + 1)$. MacNeish **(8)** in 1921 gave a construction of $n - 1$ mutually orthogonal latin squares for $n = p$ with p prime and of $n(v)$ mutually orthogonal squares of order v where

$$v = p_1^{a_1} p_2^{a_2} \ldots p_r^{a_r}$$

with $p_1, p_2, \ldots, p_r$ being distinct primes and

$$n(v) = \min(p_1^{a_1}, p_2^{a_2}, \ldots, p_r^{a_r}) - 1.$$

MacNeish conjectured that $n(v)$ was the maximum number of mutually orthogonal latin squares of order v. Both the Euler and MacNeish conjectures stood unbroken until 1959 when Parker, Shrikhande, and Bose in **(2, 3, 9, 10, 11)** showed that they were false.

While progress in the construction of mutually orthogonal latin squares was slow between 1921 and 1959, their importance grew for other reasons. Statisticians used them in the design of experiments and a striking connection between orthogonal latin squares and finite affine (and projective) plane geometries was discovered by Bose and others.

It is a trivial fact that for any n, there are at most $n - 1$ mutually orthogonal latin squares. When $n - 1$ such squares exist we say that the set of squares is complete. There is an easily established 1–1 correspondence between complete sets of orthogonal latin squares and finite affine (and hence projective) plane geometries. With a partial set of mutually orthogonal latin squares a partial affine plane can be constructed. Two types of finite projective plane are of particular interest, namely, the Desarguesian plane and the Veblen–Wedderburn plane. These can always be represented by a complete set of squares as follows. The basic square is the group addition table of an elementary abelian group and the remainder of the squares are obtained by a set of permutations of the rows in each of which the first row is kept fixed. One of the results of this paper is to give an algebraic characterization of all geometries which correspond to complete sets of squares which are obtained by permuting the rows of the addition table of an abelian group. Whether any such geometries apart from the Desarguesian and Veblen–Wedderburn planes exist is an open question.

Received May 1, 1960. Research supported in part by the United States Air Force Office of Scientific Research under Contract AF 49(638)-860.

In this paper the notion of an orthomorphism is introduced. This is a transformation which when applied to the addition table of an abelian group yields a square which is orthogonal to the original square. Criteria are obtained which enable one to say whether a given set of mutually orthogonal squares may be extended and properties are obtained which make hand computation rapid. By means of these properties the authors have obtained a set of 5 mutually orthogonal latin squares of order 12. This number exceeds the possible number given in the recent work of Parker, Shrikhande, and Bose since for n not a prime power their methods cannot yield more than $\sqrt{n}$ mutually orthogonal squares.

An algorithm suitable for machine computation has been obtained. This algorithm has been programmed by Parker and van Duren for the case $n = 12$ on the UNIVAC M 460. Exhaustive computation has shown that 5 is the maximum number of mutually orthogonal squares of order 12 obtainable by permutation of the rows of the non-cyclic abelian group of order 12. However, there are several non-isomorphic sets. Parker has also obtained the result that for $n = 15$ it is impossible to find a complete set of squares by permuting the rows of the group of order 15. This work is not yet published. As this paper is being written Bose has given the authors a report in which similar work on machine computation is being carried out at the Case Institute of Technology by two of his students.

Besides aiding in the construction of orthogonal latin squares, the theory of orthomorphisms sheds much light on finite projective planes. For instance, in the case $n = 9$ it is rapidly established that there are exactly 21 sets of 8 mutually orthogonal latin squares obtained from the elementary group of order 9, by permuting its rows. Three of the sets correspond to the Desarguesian plane, 9 to the Veblen–Wedderburn plane, and 9 to the dual of the Veblen–Wedderburn plane. The 5 possible multiplication tables of the coordinate systems are obtained as an automatic side result. One of the tables is $GF(3^2)$, the other four being the four possible Veblen–Wedderburn multiplication tables of order 9, obtained first by Marshall Hall in **(7)**.

2. Definitions and elementary properties. A latin square of order n is an n by n matrix each of whose rows and columns is a permutation of a set S of n elements. Two n by n matrices $A = (a_{ij})$ and $B = (b_{ij})$ are said to be orthogonal if the n^2 pairs (a_{ij}, b_{ij}) $(i = 1, 2, \ldots, n; j = 1, 2, \ldots, n)$ are all distinct. Note that the entries of B need not be taken from the same set as those of A. Let T be the set of elements which occur as entries of B. In this paper the authors make the convention that any latin square is orthogonal to itself, although obviously the condition of orthogonality is violated. If one considers the set of all n pairs (a_{ij}, b_{ij}) where a_{ij} is a fixed element of S, the elements b_{ij} are all the elements of T, and the set of cells (i, j) at which these b_{ij} appear, occur one in each row and one in each column of B. These entries of B are said to form a *transversal*, and B can be dissected into n mutually

exclusive transversals. Conversely, if the latin square B can be dissected into n mutually exclusive transversals, a square A orthogonal to B is obtained by assigning to all the cells of any transversal the same element of S, and assigning to different transversals different elements of S. If the entries of an n by n square A are the elements of an additive group G, with 0 in the (1,1) position, and the first row and the first column are permutations of the elements of G, then A is said to be a *group addition table* provided that the entry in the (i,j) position of A is the sum of the entries in the $(i,1)$ and $(1,j)$ positions of A. A *group addition table* is said to be in *standard form* if the entries along the main diagonal are all 0. For an abelian group G of type $a_1 \times a_2 \times a_3 \ldots \times a_r$ the standard form may even be more specialized into *computational standard form* as follows: the elements of G are taken as r-tuples $(b_1, b_2, \ldots, b_r)$ with b_i ranging from 0 to $a_i - 1$, and the first column of A is to consist of the elements of G arranged lexicographically in ascending order. For all theorems below referring to machine computations it is implied that the basic square will be in computational standard form. For a group addition table it is convenient to label the rows and columns of the square A using elements of G as labels. Any row of A will be labelled by its first entry, and the ith column of A will be given the same label as the ith row of A. Hence, if A is a group addition table in standard form and the ith column of A is given a label g, then the first entry in the ith column of A is $-g$. Each cell of A is given a double label, namely the pair (g,h) where g is the row label and h is the column label of the cell. The entry in the cell (g,h) is $g - h$ whenever the square A is in standard form.

An important folk theorem in the theory of orthogonal latin squares is based on a type of Kronecker product. Let A be a square with entries a_{ij} and for any symbol k define the square A^k as the square whose entries are the pairs (a_{ij}, k). If A and B are squares of order n and m respectively the Kronecker product square is defined as the squares $A \times B$ given by:

$$A \times B = \begin{pmatrix} A^{b_{11}}, & A^{b_{12}}, & \ldots, & A^{b_{1m}} \\ A^{b_{21}}, & A^{b_{22}}, & \ldots, & A^{b_{2m}} \\ \vdots & & & \\ A^{b_{m1}}, & A^{b_{m2}}, & \ldots, & A^{b_{mm}} \end{pmatrix}.$$

The order of $A \times B$ is nm. If A and B are *group addition tables* in *standard* or *computational standard* form of groups G and H then $A \times B$ is the group addition table of the direct sum of G and H in standard or computation standard form. (Strictly speaking this is only true if one identifies a symbol such as $((c_1, c_2, \ldots, c_r), (d_1, d_2, \ldots, d_s))$ with $(c_1, c_2, \ldots, c_r, d_1, d_2, \ldots, d_s)$. The folk theorem mentioned above reads as follows. Let $A_1, A_2, \ldots, A_r$ and $B_1, B_2, \ldots, B_r$ be two sets of mutually orthogonal squares. Then the squares $A_1 \times B_1, A_2 \times B_2, \ldots, A_r \times B_r$ are mutually orthogonal. While not explicitly formulating this theorem, MacNeish used it in his construction given in **(8)**.

3. Orthomorphisms. Let G be a group of order n written in additive form whether abelian or not, and let A be a group addition table of G in standard form, the entries in the first column of A being $0, g_2, g_3, \ldots, g_n$. A one-one mapping ϕ of G onto itself given by $\phi: x \to x\phi$ is called an orthomorphism if $x - x\phi = y - y\phi$ implies $x = y$. There is a scanty literature on mappings of groups which are equivalent to orthomorphisms. If the mapping $x \to x\phi$ is an orthomorphism, Paige and Hall in **(12)** and **(13)** call the mapping $x \to -(x\phi)$ a *complete mapping*. Their work is concerned with the question as to whether complete mappings exist in a given group. Actually, this question can be answered completely as follows. A group G admits an orthomorphism except in the case where G is of even order and its Sylow 2-subgroup is cyclic. Under the name 1-permutations, Singer in **(15)** discusses orthomorphisms of cyclic groups of odd order.

With each orthomorphism ϕ we associate the square A_ϕ which is obtained from A by permuting its rows in such a way that the first column of A_ϕ has entries $0\phi, g_2\phi, g_3\phi, \ldots, g_{n-1}\phi, g_n\phi$. The entries in the ith row of A_ϕ are

$$g_i\phi, g_i\phi - g_2, g_i\phi - g_3, \ldots, g_i\phi - g_n.$$

By convention, we will call the identity mapping I given by $I: x \to xI = x$, an orthomorphism in order to conform to a previous convention which stated that any square is orthogonal to itself.

THEOREM 1. *If ϕ is any orthomorphism the squares A and A_ϕ are orthogonal. Conversely, if A and A_1 are orthogonal where A_1 is obtained by a permutation of the rows of A then the first column of A_1 is obtained from the first column of A by an orthomorphism.*

Proof. Let a_{ij} and b_{ij} be the entries in the (i, j) cell of A and A_ϕ respectively. Consider the pairs (a_{ij}, b_{ij}), (a_{rs}, b_{rs}). It is sufficient to show that if $a_{ij} = a_{rs}$ then $b_{ij} = b_{rs}$ if and only if $i = r$ and $j = s$. $a_{uv} = g_u - g_v$ and $b_{uv} = g_u\phi - g_v$. If $a_{ij} = a_{rs}$ then $g_i - g_j = g_r - g_s$. If also $b_{ij} = b_{rs}$ then $g_i\phi - g_j = g_r\phi - g_s$. These imply $g_i\phi - g_i = g_r\phi - g_r$, and hence $g_i = g_r$ and $g_j = g_s$. Thus $i = r$ and $j = s$.

The converse part of the theorem holds since the argument is reversible.

THEOREM 2. *The squares A_ϕ and A_ψ are orthogonal if and only if $\phi^{-1}\psi$ is an orthomorphism, and this is equivalent to $x\phi - x\psi = y\phi - y\psi$ implies $x = y$.*

The proof is the same as that of Theorem 1.

We will say that the orthomorphisms ϕ and ψ are orthogonal if the corresponding squares A_ϕ and A_ψ are orthogonal. In particular, if ϕ is any orthomorphism then ϕ is orthogonal to I; also ϕ^{-1} is an orthomorphism and is orthogonal to ϕ if and only if ϕ^2 is an orthomorphism. An automorphism α of G is an orthomorphism if and only if 0 is the only element of G fixed by α.

There is a (1–1) correspondence between transversals of A and orthomorphisms of G which is obtained as follows. Let the rows and columns of A be labelled by the elements of G as given in the previous section. The entries in the cells $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$ are a transversal if and only if the mapping $b_i \to a_i$ is an orthomorphism of G, and we will say the transversal corresponds to the orthomorphism. For example, in Fig. 1, the cells marked out by square brackets are a transversal, and correspond to the orthomorphism $0 \to 1,\ 1 \to 4,\ 2 \to 2,\ 3 \to 0,\ 4 \to 6,\ 5 \to 3,\ 6 \to 5$ of the cyclic group of order 7.

0	6	5	[4]	3	2	1
[1]	0	6	5	4	3	2
2	1	[0]	6	5	4	3
3	2	1	0	6	[5]	4
4	[3]	2	1	0	6	5
5	4	3	2	1	0	[6]
6	5	4	3	[2]	1	0

FIG. 1.

Let ϕ be an orthomorphism and let g be any element of G. The mapping $\phi^{(g)}$ defined by $x\phi^{(g)} = -g + x\phi$ for all x in G is an orthomorphism. It is an easy application of Theorem 2 to show that if g is any element of G, and if A_ϕ is orthogonal to A_ψ, then $A_{\phi^{(g)}}$ is orthogonal to A_ψ. This allows us to consider only orthomorphisms ϕ such that $0\phi = 0$. Alternatively, we need only consider permutations of the rows of A which keep the first row fixed. The transversals corresponding to such orthomorphisms are precisely those which contain the entry 0 in the cell in the upper left hand corner of A. In what follows, we will assume that the orthomorphisms ϕ are of this type, that is, $0\phi = 0$.

THEOREM 3. *If a set of orthomorphisms form a group they are mutually orthogonal.*

The proof is obvious.

4. Transformation of orthomorphisms. In this section is discussed a group of mappings $O(G)$ which map orthomorphisms of G onto orthomorphisms.

Let ϕ and ψ be two orthomorphisms of G. We will say that ϕ is isomorphic to ψ if they satisfy the following conditions. Let $\phi_1, \phi_2, \ldots, \phi_r$ be the set of all orthomorphisms which are orthogonal to ϕ, and $\psi_1, \psi_2, \ldots, \psi_s$ the corresponding set for ψ. If $r = s$, and we can relabel the ψ_i in such a way that ϕ_i is orthogonal to ϕ_j if and only if ψ_i is orthogonal to ψ_j, we will say ϕ is isomorphic to ψ and write $\phi \cong \psi$.

This concept of isomorphism is too loose for some purposes but is just right if our object is to compute a maximal set of mutually orthogonal latin squares.

The group $O(G)$ we are about to define is a group of mappings of the set of all orthomorphisms onto itself in such a way that for each element λ of $O(G)$ if $\lambda: \phi \to \psi$, then $\phi \cong \psi$.

For each $g \in G$ we define an element C_g of $O(G)$ where $C_g: \phi \to \phi C_g$, ϕC_g being defined by $x(\phi C_g) = -(g\phi) + (g + x)\phi$ for all x in G. It is obvious that ϕC_g is an orthomorphism which is isomorphic to ϕ. We will call C_g a translation. Obviously $C_0 = I$, $C_g^{-1} = C_{-g}$ and $C_gC_h = C_{g+h}$. Thus the elements C_g of $O(G)$ form a sub-group, the translation subgroup of $O(G)$.

Let α be an element of the automorphism group of G. We define B_α as the mapping $B_\alpha: \phi \to \phi B_\alpha = \alpha^{-1}\phi\alpha$. It is easily verified that $\alpha^{-1}\phi\alpha$ is an orthomorphism which is isomorphic to ϕ. In the case where ϕ is also an automorphism the mapping B_α performs an inner automorphism. Easily verified are the relations $B_\alpha B_\beta = B_{\alpha\beta}$, $C_gB_\alpha = B_\alpha C_{g\alpha}$.

Finally we introduce the transformation R by $R: \phi \to \phi R = \phi^{-1}$. It is easily verified that $\phi(RC_g) = \phi(C_{g\phi^{-1}}R)$ and $RB_\alpha = B_\alpha R$. The fact that $\phi \cong \phi^{-1}$ is not immediately obvious. It is *not* in general true that if ϕ is orthogonal to ψ then ϕ^{-1} is orthogonal to ψ^{-1}. However ϕ^{-1} is orthogonal to $\phi^{-1}\psi$ by Theorem 2. The isomorphism between ϕ and ϕ^{-1} is established as follows. If $\phi_1, \phi_2, \ldots, \phi_r$ is the set of all orthomorphisms which are orthogonal to ϕ then $\phi^{-1}\phi_1, \phi^{-1}\phi_2, \ldots, \phi^{-1}\phi_r$ is the set of all orthomorphisms which are orthogonal to ϕ^{-1}, and $\phi^{-1}\phi_i$ is orthogonal to $\phi^{-1}\phi_j$ if and only if ϕ_i is orthogonal to ϕ_j.

The group $O(G)$ is now defined to be the group generated by all C_g, B_α and R.

Conjugacy of sets of orthomorphisms is now defined as follows. The set $\{I, \phi_1, \phi_2, \ldots, \phi_r\}$ is conjugate to the set $\{I, \phi_1C_g, \phi_2C_g, \ldots, \phi_rC_g\}$ under the mapping C_g. It is conjugate to the set $\{I, \phi_1B_\alpha, \phi_2B_\alpha, \ldots, \phi_rB_\alpha\}$ under the mapping B_α. With regard to the mapping R, the set $\{I = \phi_0, \phi_1, \phi_2, \ldots, \phi_r\}$ has a set of conjugates provided at least one of the ϕ_i, $i \neq 0$ is orthogonal to the remaining set of ϕ's. If ϕ_j is orthogonal to each member of the set then the set $\{\phi_j^{-1}, \phi_j^{-1}\phi_1, \phi_j^{-1}\phi_2, \ldots, I, \ldots, \phi_j^{-1}\phi_r\}$ is conjugate to the original set. It is clear that any orthogonality relationship holding amongst the orthomorphisms of one set also holds amongst the corresponding elements of a conjugate set.

With regard to a set of orthomorphisms $I, \phi_2, \phi_3, \ldots, \phi_r$ the R multiplication table is a useful concept. It is given in Fig. 2.

	I	ϕ_2	ϕ_3	$\ldots$ ϕ_r
I	I,	ϕ_2,	ϕ_3,	$\ldots, \phi_r$
ϕ_2^{-1}	ϕ_2^{-1},	I,	$\phi_2^{-1}\phi_3$,	$\ldots, \phi_2^{-1}\phi_r$
ϕ_3^{-1}	ϕ_3^{-1},	$\phi_3^{-1}\phi_2$,	I,	$\ldots, \phi_3^{-1}\phi_r$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
ϕ_r^{-1}	ϕ_r^{-1},	$\phi_r^{-1}\phi_2$,	$\phi_r^{-1}\phi_3$,	$\ldots,$ I

FIG. 2.

Each row of the table is a conjugate of the first row provided the set $I, \phi_2, \phi_3, \ldots, \phi_r$ consists of mutually orthogonal orthomorphisms. It can also be said that a necessary and sufficient condition for a set of orthomorphisms to be mutually orthogonal is that the entries of its R multiplication table are all orthomorphisms.

5. Complete sets of orthomorphisms. If a set S of $(n - 1)$ mutually orthogonal orthomorphisms of an abelian group G of order n exists, a projective geometry can be constructed. In this section it is shown how to introduce a multiplication amongst the elements of G and how to set up a corresponding analytic geometry. Let the elements of G be ordered $0, g_2, g_3, \ldots, g_n$, where $g_2, g_3, \ldots, g_n$ are an arbitrary ordering of the non-zero elements of G. We arbitrarily designate g_2 as a unit element and denote it by 1. The orthomorphism ϕ which maps $0 \to 0$ and $g_i \to g_i\phi$ will be written down as a column

$$\begin{array}{c} 1\phi \\ g_3\phi \\ g_4\phi \\ \cdot \\ \cdot \\ \cdot \\ g_n\phi \end{array}$$

Note that the element $0 = 0\phi$ is omitted from the list. If ϕ_1 and ϕ_2 are mutually orthogonal orthomorphisms, $1\phi_1 \neq 1\phi_2$, since in that case $\phi_1^{-1}\phi_2$ would map $0 \to 0$, $a \to a$ where $a = 1\phi_1$. Hence $\phi_1^{-1}\phi_2$ is not an orthomorphism, a contradiction. Hence, there are at most $n - 1$ mutually orthogonal orthomorphisms. If a full set of such orthomorphisms exist, then for each x in G there is a unique orthomorphism of the set which maps $1 \to x$. Denote this orthomorphism by ϕ_x. Hence $1\phi_x = x$. The identity orthomorphism is denoted by ϕ_1. Now form a table whose columns are $\phi_1, \phi_{g3}, \ldots, \phi_{gn}$, see Table I. The

TABLE I

ϕ_1	ϕ_{g3}	ϕ_y	ϕ_{gn}
1	g_3	y	g_n
g_3	$g_3\phi_{g3}$	.	.
g_4	$g_4\phi_{g3}$	.	.
.			
.	.	.	.
.			
x	.	$x\phi_y$	.
.		.	
.	.	.	
.		.	
g_n	$g_n\phi_{g3}$	$g_n\phi_y$	$g_n\phi_{gn}$

table may now be considered as a multiplication table the entry in any cell being the product of the entry at the extreme left in its row and the entry in the top of its column. Thus $x \cdot y = x\phi_y$ by definition. Since any two columns are orthogonal to each other the mapping of the ith column into the jth column is an orthomorphism. The relation $x\phi - x\psi = y\phi - y\psi$ implies $x = y$ can also be written $-y\phi + x\phi = -y\psi + x\psi$ implies $x = y$. This second way of writing the relation implies that the mapping of the ith row of the multiplication table into the jth row is a dual orthomorphism. By a dual orthomorphism we mean a mapping $x \to x\psi$ of G onto G which satisfies the condition $-(x\psi) + x = -(y\psi) + y$ implies $x = y$. Of course, in the case of abelian groups there is no distinction between an orthomorphism and a dual orthomorphism. Denote by ψ_x the dual orthomorphism obtained by mapping the first row into the row which starts with x. Hence $x = 1\psi_x$. Also $x \cdot y = x\phi_y = y\psi_x$ and $1x = x1 = x$. The condition that the mapping of any column into any other column is an orthomorphism is simply that the equation $x\phi_y - x\phi_z = u\phi_y - u\phi_z$ with $y \neq z$ implies $x = u$. Hence $xy - xz = uy - uz$ implies $x = u$ or $y = z$. This can be stated in the alternative form namely: the equation $xa = c + xb$ has a unique solution if $a \neq b$, and this is equivalent to the statement $ay = by + c$ has a unique solution provided $a \neq b$. Conversely, let multiplication be introduced in G in an arbitrary way subject only to the conditions $xa = c + xb$ has a unique solution whenever $a \neq b$ and $x0 = 0$. Consider the set $a0 - b0 = 0$, $ag_2 - bg_2$, $ag_3 - bg_3$, $\ldots$, $ag_n - bg_n$. If these are all distinct it implies that the equation $ay = by + c$ has a unique solution. If on the other hand $ag_i - bg_i = ag_j - bg_j$ for $i \neq j$, then $ag_i = ag_j + (-bg_j + bg_i)$. Hence the equation $xg_i = xg_j + (-bg_j + bg_i)$ has two solutions namely $x = a$ and $x = b$, a contradiction. Thus for a finite group G, any introduced system of multiplication satisfying the conditions $x0 = 0$ and $xa = c + xb$ with $a \neq b$ has a unique solution also satisfies the condition $ay = by + c$ has a unique solution. Also the mapping $xa \to xb$ where a and b are fixed and x ranges over G is an orthomorphism so that the columns of the multiplication table form a complete set of mutually orthogonal orthomorphisms.

An analytic geometry can now be introduced. We assume that G has a unit element under multiplication and the equation $xa = xb + c$ has a unique solution if $a \neq b$. For the points of the geometry we take the triplets $(a, b, 1)$, $(a, 1, 0)$, and $(1, 0, 0)$. For the lines we take the equations $x + Ay + Bz = 0$, $y + Bz = 0$, $z = 0$. It is readily verified that the points and lines form a projective plane. At present the only known finite planes of this type are the Desarguesian plane and the Veblen–Wedderburn plane.

We now interpret the distributive laws of multiplication. The left distributive law $x \cdot (y + z) = xy + xz$ becomes $x\phi_{y+z} = x\phi_y + x\phi_z$, which says that the sum of two columns of the multiplication table is a third column. Alternatively this law may be written $(y + z)\psi_x = y\psi_x + z\psi_x$, which shows that the mapping ψ_x is an automorphism. Hence, a left distributive law is equivalent to the

condition that the mapping of the first row into any other row is an automorphism. Similarly, the right distributive law is equivalent to the statement that the mapping of the first column into any other column is an automorphism.

Conjugacy takes on some interesting properties here in the case where G is abelian. In general, if a complete set of orthomorphisms is replaced by a conjugate set under the group $O(G)$, then the multiplication table for the second set is left (right) distributive if and only if the same holds for the first set. We prove it for the case of conjugacy under C_g only. Let $\phi_1, \phi_{g_2}, \phi_{g_4}, \ldots, \phi_{g_{n-1}}$ be a complete set of orthomorphisms for which the left distributive law holds. This means that $x\phi_y + x\phi_z = x\phi_{y+z}$ and also that the mapping $\Lambda(x, y): x\phi_z \to y\phi_z$, where z ranges over G, is an automorphism for each x, y in G. It is sufficient to show that $x(\phi_z C_g) \to y(\phi_z C_g)$ where z ranges over G with x, y, g fixed is an automorphism. Now

$$\begin{aligned}
x(\phi_z C_g) + x(\phi_u C_g) &= -(g\phi_z) + (g + x)\phi_z - g\phi_u + (g + x)\phi_u \\
&= (g + x)\phi_z + (g + x)\phi_u - (g\phi_z + g\phi_u) \\
&= (g + x)\phi_{z+u} - g\phi_{z+u} \\
&= x(\phi_{z+u}C_g) \to y(\phi_{z+u}C_g) \\
&= -(g\phi_{z+u}) + (g + y)\phi_{z+u} \\
&= -(g\phi_z) - (g\phi_u) + (g + y)\phi_z + (g + y)\phi_u \\
&= y(\phi_z C_g) + y(\phi_u C_g)
\end{aligned}$$

as required. For non-abelian groups, the distributive law may not be invariant under conjugacy.

The results of this section are summed up as follows:

THEOREM 4. *Let A be the group addition table of a group G. A necessary and sufficient condition that a complete set of orthogonal latin squares obtainable from A by permutation of its rows exist is that it is possible to define a multiplication in G such that $0x = x0 = 0$ and such that the equation $xa = c + xb$ has a unique solution in G provided $a \neq b$. If G is abelian and the multiplication satisfies a left (right) distributive law, then so does the multiplication obtainable from a conjugate set of orthomorphisms.*

6. A machine computation algorithm. The theory of orthomorphisms leads very readily to an algorithm for the computation of orthogonal latin squares, which is easy to program on a digital computer, and which takes a relatively short time to compute. We quote the result without proof. Let A be a group addition table in computational standard form of a group G. Let $I, \phi_2, \ldots, \phi_r$ be a set of mutually orthogonal orthomorphisms and $A, A_{\phi_2}, \ldots,$

A_{ϕ_r} the corresponding squares. This set of squares *except for* A is transposed into the set $A, A_{\phi_2}^T, A_{\phi_3}^T, \ldots, A_{\phi_r}^T$. A necessary and sufficient condition that a latin square exist and be orthogonal to $A, A_{\phi_2}, \ldots, A_{\phi_r}$ is that the transposed set of squares, together with A, have a common transversal passing through the cell in the upper left corner. The orthomorphism ϕ_{r+1} corresponding to this transversal is orthogonal to all preceding orthomorphisms. Some actual machine results will be quoted later on, but a systematic report on machine computation will appear in a subsequent paper.

7. Analysis of some cases. As applications of the previous theory some examples of systems of orthogonal latin squares for small n will be given. Throughout this section we will use the symbol $\{a\} \times \{b\} \times \ldots \times \{r\}$ to denote the direct product of cyclic groups of orders $a, b, \ldots, r$. No examples of orthomorphisms of non-abelian groups are given here. The dihedral groups of orders 8 and 12, as well as the alternating group of order 12, are of interest, but our analysis is not yet complete.

For $n = 3$ or 5, complete systems of squares are obtained, and these correspond to automorphisms of $\{3\}$ and $\{5\}$. For $n = 4$, the group $\{4\}$ has no orthomorphisms while the group $\{2\} \times \{2\}$ has exactly 3, these being a complete set. The automorphism group of $\{2\} \times \{2\}$ is S_3 and the elements of A_3 are all the orthomorphisms. For $n = 6$, the group $\{6\}$ has no orthomorphisms.

The case $n = 7$ is the first value of n for which orthomorphisms which are not automorphisms exist. There is a complete set of 6 mutually orthogonal orthomorphisms corresponding to the automorphism group of $\{7\}$, together with a set of 14 maverick orthomorphisms each of which is orthogonal only to itself and the identity. This set of 14 is a complete set of conjugates of any one of them under the group $O(\{7\})$. Denoting by $\{a_0 a_1 \ldots a_6\}$ the orthomorphism $i \to a_i$ the list is as follows, the first 6 being automorphisms:

$$
\begin{array}{lll}
\{0\ 1\ 2\ 3\ 4\ 5\ 6\}, & \{0\ 2\ 4\ 6\ 1\ 3\ 5\}, & \{0\ 3\ 6\ 2\ 5\ 1\ 4\} \\
\{0\ 4\ 1\ 5\ 2\ 6\ 3\}, & \{0\ 5\ 3\ 1\ 6\ 4\ 2\}, & \{0\ 6\ 5\ 4\ 3\ 2\ 1\} \\
\{0\ 3\ 1\ 6\ 5\ 2\ 4\}, & \{0\ 2\ 5\ 1\ 6\ 4\ 3\}, & \{0\ 3\ 6\ 4\ 2\ 1\ 5\} \\
\{0\ 4\ 6\ 2\ 5\ 3\ 1\}, & \{0\ 6\ 3\ 5\ 1\ 4\ 2\}, & \{0\ 5\ 4\ 1\ 3\ 6\ 2\} \\
\{0\ 5\ 3\ 2\ 6\ 1\ 4\}, & \{0\ 3\ 5\ 2\ 1\ 6\ 4\}, & \{0\ 3\ 6\ 1\ 5\ 4\ 2\} \\
\{0\ 5\ 3\ 6\ 2\ 4\ 1\}, & \{0\ 6\ 4\ 2\ 5\ 1\ 3\}, & \{0\ 4\ 3\ 1\ 6\ 2\ 5\} \\
\{0\ 5\ 1\ 4\ 6\ 3\ 2\}, & \{0\ 2\ 6\ 5\ 3\ 1\ 4\}. &
\end{array}
$$

The case $m = 8$. The group $\{8\}$ has no orthomorphisms. The group $\{4\} \times \{2\}$ has no orthomorphisms which are automorphisms, but has 49 orthomorphisms. These separate into 24 sets of 3 mutually orthogonal orthomorphisms, the identity being included in each set. Each triplet is conjugate

to any other triplet under $O(\{4\} \times \{2\})$. They are listed below as pairs with the identity omitted. The elements of $\{4\} \times \{2\}$ will be denoted by 0, 1, 2, 3, $0'$, $1'$, $2'$, $3'$ and $\{a_0\, a_1\, a_2\, a_3\, a_0'\, a_1'\, a_2'\, a_3'\}$ will denote the orthomorphism $i \to a_i$, $i' \to a_i'$

$\{0\ 2'\ 3\ 1'\ 3'\ 1\ 0'\ 2\}$,	$\{0\ 3\ 3'\ 2'\ 1'\ 0'\ 2\ 1\}$;
$\{0\ 0'\ 3\ 3'\ 2'\ 2\ 1'\ 1\}$,	$\{0\ 3\ 1'\ 0'\ 2\ 1\ 3'\ 2'\}$;
$\{0\ 2\ 3'\ 1'\ 2'\ 0'\ 1\ 3\}$,	$\{0\ 1'\ 1\ 2'\ 2\ 3'\ 3\ 0'\}$;
$\{0\ 1'\ 3'\ 2\ 3\ 2'\ 0'\ 1\}$;	$\{0\ 0'\ 3\ 1'\ 1\ 3'\ 2\ 2'\}$;
$\{0\ 3'\ 1\ 2'\ 1'\ 2\ 0'\ 3\}$,	$\{0\ 2'\ 1'\ 1\ 3'\ 3\ 2\ 0'\}$;
$\{0\ 1'\ 1\ 0'\ 2'\ 3\ 3'\ 2\}$,	$\{0\ 0'\ 3'\ 1\ 2\ 2'\ 1'\ 3\}$;
$\{0\ 3'\ 1'\ 2\ 2'\ 1\ 3\ 0'\}$,	$\{0\ 2'\ 3\ 3'\ 2\ 0'\ 1\ 1'\}$;
$\{0\ 2\ 1'\ 3'\ 1\ 3\ 0'\ 2'\}$,	$\{0\ 3'\ 1\ 0'\ 3\ 2'\ 2\ 1'\}$;
$\{0\ 3'\ 2'\ 1\ 3\ 2\ 1'\ 0'\}$,	$\{0\ 2'\ 0'\ 2\ 1'\ 3'\ 1\ 3\}$;
$\{0\ 1'\ 0'\ 1\ 3'\ 2'\ 3\ 2\}$,	$\{0\ 0'\ 2'\ 2\ 1\ 3\ 3'\ 1'\}$;
$\{0\ 2'\ 0'\ 2\ 3\ 1\ 3'\ 1'\}$,	$\{0\ 3'\ 2'\ 1\ 1'\ 0'\ 3\ 2\}$;
$\{0\ 2\ 2'\ 0'\ 3'\ 3\ 1\ 1'\}$,	$\{0\ 3\ 0'\ 3'\ 1\ 2'\ 1'\ 2\}$;
$\{0\ 3\ 2'\ 1'\ 1\ 0'\ 3'\ 2\}$,	$\{0\ 2\ 0'\ 2'\ 3'\ 1\ 3\ 1'\}$;
$\{0\ 3\ 0'\ 3'\ 1'\ 2\ 1\ 2'\}$,	$\{0\ 2\ 2'\ 0'\ 3\ 3'\ 1'\ 1\}$;
$\{0\ 2\ 0'\ 2'\ 1\ 3'\ 1'\ 3\}$,	$\{0\ 3\ 2'\ 1'\ 3'\ 2\ 1\ 0'\}$;
$\{0\ 0'\ 2'\ 2\ 1'\ 3'\ 3\ 1\}$,	$\{0\ 1'\ 0'\ 1\ 3\ 2\ 3'\ 2'\}$;
$\{0\ 2'\ 1\ 3'\ 1'\ 3\ 0'\ 2\}$,	$\{0\ 3\ 1'\ 0'\ 3'\ 2'\ 2\ 1\}$;
$\{0\ 0'\ 1\ 1'\ 2'\ 2\ 3'\ 3\}$,	$\{0\ 3\ 3'\ 2'\ 2\ 1\ 1'\ 0'\}$;
$\{0\ 2\ 1'\ 3'\ 2'\ 0'\ 3\ 1\}$,	$\{0\ 1'\ 3\ 0'\ 2\ 3'\ 1\ 2\}$;
$\{0\ 3'\ 1'\ 2\ 1\ 2'\ 0'\ 3\}$,	$\{0\ 2'\ 1\ 1'\ 3\ 3'\ 2\ 0'\}$;
$\{0\ 1'\ 3\ 2'\ 3'\ 2\ 0'\ 1\}$,	$\{0\ 0'\ 3'\ 1\ 1'\ 3\ 2\ 2'\}$;
$\{0\ 3'\ 3\ 0'\ 2'\ 1\ 1'\ 2\}$,	$\{0\ 2'\ 1'\ 1\ 2\ 0'\ 3'\ 3\}$;
$\{0\ 1'\ 3'\ 2\ 2'\ 3\ 1\ 0'\}$,	$\{0\ 0'\ 1\ 3'\ 2\ 2'\ 3\ 1'\}$;
$\{0\ 2\ 3'\ 1'\ 3\ 1\ 0'\ 2'\}$,	$\{0\ 3'\ 3\ 2'\ 1\ 0'\ 2\ 1'\}$.

The group $\{2\} \times \{2\} \times \{2\}$ is the most interesting case of $n = 8$. No orthomorphisms which are not automorphisms exist. However, the automorphism group of $\{2\} \times \{2\} \times \{2\}$ is the simple group of order 168. By Sylow's theorem, there are 8 subgroups of order 7, and each of these subgroups consists of elements which are orthomorphisms. Hence there are 8 complete sets of mutually orthogonal latin squares all of which are conjugate under the group generated by the B_a. They all correspond to the Desarguesian plane of order 8.

The case $n = 9$. For the group $\{9\}$ it is easily established that a complete set does not exist. An exhaustive classification can be readily carried out, and this leads to a totality of 226 orthomorphisms. It appears that no set of 3 mutually orthogonal latin squares exists, but the calculation has not been checked.

For the group $\{3\} \times \{3\}$ the results are extremely interesting. Represent the elements of this group by 0, 1, 2, $0'$, $1'$, $2'$, $0''$, $1''$, $2''$ with addition being

mod 3 with respect to both the integers and the superscripts. The automorphism group of $\{3\} \times \{3\}$ is of order 48. Of these automorphisms, 28 are orthomorphisms. These may be designated as I, α, α^2, α^3, α^4, α^5, α^6, α^7, β, β^2, β^3, β^5, β^6, β^7, γ, γ^2, γ^3, γ^5, γ^6, γ^7, A, B, C, D, E, F, G, H, where

$$\alpha = \begin{pmatrix} 1 \to 2' \\ 0' \to 1 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 \to 0' \\ 0' \to 1'' \end{pmatrix}, \quad \gamma = \begin{pmatrix} 1 \to 1' \\ 0' \to 2' \end{pmatrix}$$

$$A = \begin{pmatrix} 1 \to 2 \\ 0' \to 1'' \end{pmatrix}, \quad B = \begin{pmatrix} 1 \to 2 \\ 0' \to 2'' \end{pmatrix}, \quad C = \begin{pmatrix} 1 \to 2' \\ 0' \to 0'' \end{pmatrix}$$

$$D = \begin{pmatrix} 1 \to 0'' \\ 0' \to 1' \end{pmatrix}, \quad E = \begin{pmatrix} 1 \to 2'' \\ 0' \to 0'' \end{pmatrix}, \quad F = \begin{pmatrix} 1 \to 0' \\ 0' \to 2' \end{pmatrix}$$

$$G = \begin{pmatrix} 1 \to 1' \\ 0' \to 2 \end{pmatrix}, \quad H = \begin{pmatrix} 1 \to 1'' \\ 0' \to 1 \end{pmatrix}.$$

There are four groups of orthomorphic automorphisms of order 8 as follows: each of α, β, γ generate a cyclic group of order 8, and the even powers of α, β, γ are the quaternion group. It is impossible to realize by a set of orthomorphisms the other possible groups of order eight, namely, the groups $\{2\} \times \{2\} \times \{2\}$, $\{4\} \times \{2\}$, and the dihedral group, since it can be readily calculated that there are exactly three orthomorphisms of order 2, no two of which are orthogonal. The cyclic groups correspond to the Desarguesian plane, and the quaternion group to the Veblen–Wedderburn plane. If the automorphisms corresponding to the quaternion group are written as rows, the columns of the table are orthomorphisms which are not automorphisms. Applying successively the transformations C_1, C_2, $C_{0'}$, $C_{1'}$, $C_{2'}$, $C_{0''}$, $C_{1''}$, $C_{2''}$ to the columns of the table, one obtains 8 other complete sets of orthomorphisms. In the resultant tables the rows represent complete sets of automorphisms. The 12 complete sets of 8 mutually orthogonal orthomorphisms are as follows:

(1)	$\{I, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6, \alpha^7\}$
(2)	$\{I, \beta, \beta^2, \beta^3, \beta^4, \beta^5, \beta^6, \beta^7\}$
(3)	$\{I, \gamma, \gamma^2, \gamma^3, \gamma^4, \gamma^5, \gamma^6, \gamma^7\}$
(4)	$\{I, \alpha^2, \beta^2, \gamma^2, \alpha^4 = \beta^4 = \gamma^4, \alpha^6, \beta^6, \gamma^6\}$
(5)	$\{I, \alpha^4, \alpha^7, \beta^7, \gamma^7, \alpha^5, \beta^5, \gamma^5\}$
(6)	$\{I, \alpha^4, \alpha, \beta, \gamma, \alpha^3, \beta^3, \gamma^3\}$
(7)	$\{I, \beta, \beta^5, \beta^6, B, E, G, H\}$
(8)	$\{I, \gamma, \gamma^5, \gamma^6, B, C, D, H\}$
(9)	$\{I, \alpha, \alpha^5, \alpha^6, B, D, E, F\}$
(10)	$\{I, \alpha^2, \alpha^3, \alpha^7, A, C, G, H\}$
(11)	$\{I, \gamma^2, \gamma^3, \gamma^7, A, E, F, G\}$
(12)	$\{I, \beta^2, \beta^5, \beta^7, A, C, D, F\}$

If the multiplication tables in cases 4 to 12 are transposed one obtains 9 further sets of mutually orthogonal orthomorphisms. It is interesting to note

what the multiplication tables are in the various cases. In cases (1), (2), (3) the table is $GF(3^2)$. In case (4) it is the near field of order 9. In case (5) it is the Veblen–Wedderburn–Hall system with equation $x^2 = x + 1$. In case (6) it is the Veblen–Wedderburn–Hall system with equation $x^2 = 2x + 1$ and in cases (7) to (12) it is the Veblen–Wedderburn system with 2 not in the centre. These multiplication tables were first obtained by Hall **(7)**, from an entirely different viewpoint.

The 21 sets of 8 mutually orthogonal latin squares are all that there are. The following R multiplication table shows that cases (5) to (12) are conjugate under R.

	I	α^4	α	α^3	β	β^3	γ	γ^3
I	I	α^4	α	α^3	β	β^3	γ	γ^3
α^4	α^4	I	α^5	α^7	β^5	β^7	γ^5	γ^7
α^7	α^7	α^3	I	α^2	H	A	C	G
α^5	α^5	α	α^6	I	E	D	B	F
β^7	β^7	β^3	F	C	I	β^2	D	A
β^5	β^5	β	B	G	β^6	I	H	E
γ^7	γ^7	γ^3	E	A	G	F	I	γ^2
γ^5	γ^5	γ	D	H	B	C	γ^6	I

There are many orthomorphisms of $\{3\} \times \{3\}$ which are not part of a complete set. These will be reported in a subsequent paper.

The cases $n = 10$ and $n = 11$. The group $\{10\}$ has no orthomorphisms, while the group $\{11\}$ has a complete set together with several maverick orthomorphisms as in the case $n = 7$. Calculations of these maverick orthomorphisms lead to a totality 3432, exclusive of the automorphisms.

The case $n = 12$. The group $\{12\}$ has no orthomorphisms. The group $\{6\} \times \{2\}$ has, besides the identity, only two other orthomorphic automorphisms, and to this pair there does not exist an orthomorphism which is orthogonal to both. Some principles of construction will now be stated and criteria which enable one to determine when a set of orthogonal orthomorphisms cannot be extended will be given. The results carry over completely for the case $n = 4(2k + 1)$, and similar methods can be established for other n.

The elements of the group $\{6\} \times \{2\}$ will be denoted by $0, 1, 2, 3, 4, 5, 0', 1', 2', 3', 4', 5'$, with rules of addition $a + b' = (a + b)'$ and $a' + b' = a + b$, where addition is mod 6. $\{6\} \times \{2\}$ has three subgroups of order 6, namely,

(1) $0, 1, 2, 3, 4, 5$
(2) $0, 1', 2, 3', 4, 5'$
(3) $0, 2', 4, 0', 2, 4'$;

one subgroup of order 4, namely, $0, 3, 0', 3'$; one subgroup of order 3, namely, $0, 2, 4$; and three subgroups of order 2, namely, $0, 3$; $0, 0'$; $0, 3'$.

The computational standard form is given by the square

0	5	4	3	2	1	0′	5′	4′	3′	2′	1′
1	0	5	4	3	2	1′	0′	5′	4′	3′	2′
2	1	0	5	4	3	2′	1′	0′	5′	4′	3′
3	2	1	0	5	4	3′	2′	1′	0′	5′	4′
4	3	2	1	0	5	4′	3′	2′	1′	0′	5′
5	4	3	2	1	0	5′	4′	3′	2′	1′	0′
0′	5′	4′	3′	2′	1′	0	5	4	3	2	1
1′	0′	5′	4′	3′	2′	1	0	5	4	3	2
2′	1′	0′	5′	4′	3′	2	1	0	5	4	3
3′	2′	1′	0′	5′	4′	3	2	1	0	5	4
4′	3′	2′	1′	0′	5′	4	3	2	1	0	5
5′	4′	3′	2′	1′	0′	5	4	3	2	1	0

The four blocks of the square will be denoted by I, II, III, IIII according to the pattern

I	II
III	IIII

Since there is a one-one correspondence between transversals and orthomorphisms, these terms will be used interchangeably throughout. Several properties of transversals will now be stated.

Two transversals are said to agree in a column if the cells belonging to each one in the column are in the same block. Two transversals are said to have agreement of type $[r, s]$ if they have r agreements in columns of blocks I and III and s agreements in columns of blocks II and IIII. The division of the square into four blocks is really a division with respect to the subgroup 0, 1, 2, 3, 4, 5. With regard to the two other subgroups of order 6 a similar subdivision may be effected. Also, the notion of $[r, s]$ agreement, here defined, is really a concept associated with the subgroup 0, 1, 2, 3, 4, 5. We can define an $[r, s]$ agreement modulo each of the remaining subgroups of order 6.

PROPERTY 1. *For any transversal each of the blocks* I, II, III, IIII *contains three cells.* (*This is also true for the division of the square into blocks with respect to each of the other two subgroups.*)

PROPERTY 2. *If two transversals have* $[r, s]$ *agreement then* r *and* s *are both even.*

PROPERTY 3. *If two transversals have* $[r, s]$ *agreement and are orthogonal then* $r + s = 6$.

Properties 1, 2, 3 are easily established and will not be proved here. From Property 3, it follows that two mutually orthogonal transversals have agreement of type $[6, 0]$, $[4, 2]$, or $[2, 4]$. (Agreement of type $[0, 6]$ is excluded

since we are considering only transversals through the cell in the upper left corner of the square.)

PROPERTY 4. *If two transversals have* [6, 0] *agreement modulo any one of the three subgroups of order* 6, *there does not exist a transversal mutually orthogonal to both.*

Proof. Denote by α and β the two transversals with [6, 0] agreement. If γ is any transversal having $[r, s]$ agreement with α, then γ has $[r, 6 - s]$ agreement with β. If γ is orthogonal to both α and β then $r + s = 6$ and $r + (6 - s) = 6$. Hence, $r = s = 3$, which contradicts Property 2.

It follows that if a set of mutually orthogonal transversals contains at least 3, then every pair has either [4, 2] or [2, 4] agreement modulo each of the subgroups, of order 6.

With the above properties alone, hand computation has yielded 5 mutually orthogonal latin squares of order 12.

The above properties are essentially modulo 2 properties. It is possible to give modulo 3 properties but these are omitted as they did not aid in the computations.

We state some computed results. As before the orthomorphism $i \rightarrow a_i$, $i' \rightarrow a_i'$ will be written $\{a_0, a_1, a_2, a_3, a_4, a_5; a_0', a_1', a_2', a_3', a_4', a_5'\}$. The identity orthomorphism has been omitted from all lists.

Examples.

(1) The transversals

$$\{0\ 0'\ 2'\ 2\ 1'\ 1;\quad 3'\ 5'\ 4\ 4'\ 5\ 3\}$$
$$\{0\ 4'\ 1'\ 5\ 2'\ 3;\quad 2\ 1\ 5'\ 4\ 3'\ 0'\}$$

are orthogonal and have [6, 0] agreement modulo the subgroup 0, 1, 2, 3, 4, 5.

(2) The transversals

$$\{0\ 0'\ 2'\ 2\ 1'\ 1;\quad 3'\ 5'\ 4\ 4'\ 5\ 3\}$$
$$\{0\ 5'\ 4'\ 2'\ 2\ 4;\quad 1\ 3'\ 5\ 3\ 1'\ 0'\}$$

are orthogonal and have [6, 0] agreement modulo the subgroup 0, 1′, 2, 3′, 4, 5′.

(3) The transversals

$$\{0\ 0'\ 2'\ 2\ 1'\ 1;\quad 3'\ 5'\ 4\ 4'\ 5\ 3\}$$
$$\{0\ 5'\ 4'\ 5\ 1\ 2';\quad 1'\ 2\ 0'\ 3\ 3'\ 4\}$$

are orthogonal and have [6, 0] agreement modulo the subgroup 0, 2′, 4, 0′, 2, 4′.

(4) The four transversals

$$\{0\ 0'\ 2'\ 2\ 1'\ 1;\quad 3'\ 5'\ 4\ 4'\ 5\ 3\}$$
$$\{0\ 3\ 0'\ 1\ 3'\ 5';\quad 2\ 2'\ 5\ 4\ 1'\ 4'\}$$
$$\{0\ 2'\ 1\ 5'\ 5\ 3';\quad 3\ 4'\ 2\ 1'\ 0'\ 4\}$$
$$\{0\ 4\ 5'\ 4'\ 2\ 1';\quad 2'\ 0'\ 3'\ 1\ 3\ 5\}$$

are mutually orthogonal and together with the identity yield 5 mutually orthogonal latin squares of order 12.

The algorithm given in § 6 has been programmed by Parker and van Duren for the UNIVAC M–460. Many sets of 5 mutually orthogonal latin squares exist but no set of six. There exist transversals with as many as 48 transversals orthogonal to them. An example of one such transversal is $\{0\,4'\,4\,2'\,2\,0';\ 5'\,5\,3'\,3\,1'\,1\}$. There also exist configurations consisting of four sets of 5 mutually orthogonal latin squares with three of the squares common to all four sets. Apart from the identity there are exactly 16,512 orthomorphisms and apart from isomorphism there are exactly four sets of 5 mutually orthogonal latin squares.

A detailed analysis of the non-isomorphic cases will appear in a subsequent paper.

8. Concluding remarks. The problem of finding a complete set of squares for n not a prime power is still open. However, even in the case where n is a prime power there is a possibility of discovering planes of a new type. If Veblen–Wedderburn planes were the only type obtainable from orthomorphisms of a group, this would imply that a finite system which was a group under addition, and which had a multiplication for which the equation $ax = bx + c$ had a unique solution if $a \neq b$, would of necessity satisfy at least one distributive law. This does not seem likely. In the infinite case, there are planes not of the Veblen–Wedderburn type which belong to such a system. An example is given in Pickert **(14)**.

For $n = 4p$, with p an odd prime it is conjectured that using orthomorphisms at least $2p - 1$ mutually orthogonal latin squares can be constructed. A complete set is not ruled out. It may be noted that for $n = 4p$, a pair of squares with $[2p, 0]$ agreement does not have a third square orthogonal to it. For $n = 8p$, no such criterion exists. Perhaps the search for a complete set of squares should be sought in these values of n. The smallest is $n = 24$ and this is just on the verge of impracticality for machine computation.

What appears to be the biggest lack is a positive construction for orthomorphisms which are not automorphisms. Bruck **(4)** has shown that using automorphisms only the MacNeish estimate cannot be exceeded. Our present results enable a rapid calculation of orthomorphisms by giving a number of criteria which enable us to reject cases early in the computation. For large n, these criteria are not enough and the calculation is impractical.

References

1. R. C. Bose and K. R. Nair, *On complete sets of Latin squares*, Sankhya, *5* (1941), 361–382.
2. R. C. Bose and S. S. Shrikhande, *On the falsity of Euler's conjecture about the non-existence of two orthogonal Latin squares of order 4t+2*, Proc. Nat. Acad. Sci. U.S.A., *45* (1959), 734–737.

3. R. C. Bose, S. S. Shrikhande, and E. T. Parker, *Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture*, Can. J. Math., *12* (1960), 189–203.
4. R. H. Bruck, *Finite nets I, numerical invariants*, Can. J. Math., *3* (1951), 94–107.
5. A. L. Dulmage, D. M. Johnson, and N. S. Mendelsohn, *Orthogonal Latin squares*, Can. Math. Bull., *2* (1959), 211–216.
6. L. Euler, *Recherches sur une nouvelle espece des quarrés magiques*, Verh. Zeeuwsch Genoot. Weten Vliss, *9* (1782), 85–239.
7. Marshall Hall, *Projective planes*, Trans. Amer. Math. Soc., *54* (1943), 229–277.
8. H. F. MacNeish, *Euler squares*, Ann. Math., *23* (1921), 221–227.
9. E. T. Parker, *Construction of some sets of pairwise orthogonal latin squares*, Notices Amer. Math. Soc., *5* (1958), 815 (Abstract).
10. ——— *Construction of some sets of mutually orthogonal latin squares*, Proc. Amer. Math. Soc., *10* (1959), 949–951.
11. ——— *Orthogonal Latin squares*, Proc. Nat. Acad. Sci. U.S.A., *45* (1959), 859–862.
12. L. J. Paige, *Complete mappings of finite groups*, Pac. J. Math., *1* (1951), 111–116.
13. L. J. Paige and Marshall Hall, *Complete mappings of finite groups*, Pac. J. Math., *5* (1955), 541–549.
14. Gunter Pickert, *Projective Ebene* (Springer, Berlin), page 90.
15. J. Singer, *A class of groups associated with Latin squares*, Amer. Math. Monthly, *67* (1960), 235–240.
16. G. Tarry, *Le problème des 36 officiers*, Ass. France Av. Sci., *29* (1900), 170–203.

University of Manitoba

PARTITION RINGS OF CYCLIC GROUPS OF ODD PRIME POWER ORDER[1]

K. I. APPEL

A ring R over a commutative ring K, that has a basis of elements $g_1, g_2, \ldots, g_n$ forming a group G under multiplication, is called a *group ring* of G over K. Since all group rings of a given G over a given K are isomorphic, we may speak of the group ring KG of G over K.

Let π be any partition of G into non-empty sets $G_A, G_B, \ldots$. Any subring P of KG that has a basis of elements

$$A = \sum_{g \in G_A} m_g g, \ldots, m_g \neq 0 \text{ in } K,$$

is a *partition ring* of G over K.

If P is a partition ring of G over Z, the ring of integers, then the basis $A, B, \ldots$ for P clearly serves as a basis for a partition ring $P' = Q \otimes P$ of G over Q, the field of rationals. If, in addition, for each $A, B, \ldots$ the coefficients m_g, all $g \in G_A$, have no common factor, we shall call $A, B, \ldots$ a *reduced integral basis* for P'.

LEMMA. *Every partition ring P over the rationals has a reduced integral basis.*

By hypothesis, the ring P has a basis of elements

$$A = \sum_{g \in G_A} (u_g/v_g)g$$

where u_g, v_g are non-zero integers. We can write $A = (u_A/v_A)\Sigma m_g g$ where the $m_g \neq 0$ are integers without any common factor. Then the $A' = \Sigma m_g g$ forms a basis for P, and it remains to show that in the multiplication table,

$$A'_i A'_j = \sum b_{ij}{}^k A'_k,$$

the rationals $b_{ij}{}^k$ are in fact all integers. Fix i, j, and k, and consider $g \in G_{A_k}$. Since all coefficients on the left are clearly integers, the same is true on the right, and $b_{ij}{}^k m_g$ is an integer for each $g \in G_{A_k}$. Since the m_g have no common factor, this requires that $b_{ij}{}^k$ be an integer.

Henceforth, by partition ring we mean integral partition ring over the rationals, and by basis, we mean reduced integral basis, We will also adopt the convention that basis elements be chosen such that for each G_A at least one m_g is positive.

Received September 1, 1959; in revised form March 24, 1960.

[1]This paper is based on a portion of a dissertation submitted to the Graduate School of the University of Michigan in partial fulfilment of the requirements for the Ph.D. degree. The author wishes to thank Professor R. C. Lyndon for his advice and encouragement.

Let G be a finite abelian group. For each integer y prime to the order of G, define (y) to be the map $g \to g^y$ for each g in G. Then (y) is an automorphism of G, and we will call the automorphisms of this type the power automorphisms of G.

THEOREM 1. *Let G be a finite abelian group, and y an integer prime to the order n of G. Let A be a basis element of a partition ring P of G. Then there exists an element B, of the same basis, such that*

$$A^{(y)} = \sum_{g \in G_A} m_g g^y = \pm B.$$

Proof. First, we show that $G_A^{(y)} = \{g^y | g \in G_A\}$ is a union of partition classes under the partition induced by P. Assume not. Then there exists a basis element $D = ag_0^y + bg_1 + \ldots$ where $g_0 \in G_A$, $g_1 \neq g^y$ for any g in G_A. (Here we use $\ldots$ in a special sense meaning that a and b are the full coefficients of g_0^y and g_1 respectively, that is, the elements occurring in the remaining terms of the sum are distinct from g_0^y and g_1. In similar contexts the same convention is employed.)

Now we employ the theorem of Dirichlet that if j and k are relatively prime there exist infinitely many primes congruent to j modulo k. Since $(y, n) = 1$, by Dirichlet's theorem, we may choose $q \equiv y \pmod{n}$ such that $q > |m_g|$, all $g \in G_A$, $q > |b|$, q prime. But, modulo q,

$$A^q \equiv A^{(q)} = A^{(y)} = \sum_{g \in G_A} m_g g^y.$$

Since (y) is an automorphism of G, $g_1 \neq g_2$ implies $g_1^y \neq g_2^y$.

A^q must be a sum of basis elements of P. Therefore $A^q = ug_0^y + \ldots$, $A^q = kD + \ldots = kag_0^y + kbg_1 + \ldots$. However, $g_1 \neq g^y$ for any g in G_A so $q | kb$, and since $q > |b|$, $q | k$. But

$$ka \equiv m_{g_0} \pmod{q}$$

and

$$|m_{g_0}| < q,\ m_{g_0} \neq 0 \text{ so } q \nmid ka,$$

which is a contradiction.

Next, we show that $G_A^{(y)}$ is a partition set of the partition induced by P. Suppose not. Let $G_A^{(y)} = G_B \cup G_C \cup \ldots$. Let $yz \equiv 1 \pmod{n}$. Then $G_B^{(z)} \subset G_A^{(yz)} = G_A$. Since by the above, $g_B^{(z)}$ is a union of partition classes while G_A is a single partition class this implies that $G_B^{(z)} = G_A$ and $G_A^{(y)} = G_B$.

Write

$$B = \sum_{g \in G_A} n_g g^y$$

for the basis element corresponding to the partition class $G_A^{(y)}$. If q is a prime with $q \equiv y \pmod{n}$,

$$A^q = \Big(\sum_{g \in G_A} m_g g\Big)^q \equiv \sum_{g \in G_A} m_g g^y \pmod{q}.$$

It follows that B appears in the product A^q with non-zero coefficient, say

$$A^q = \lambda(q)B + \ldots = \sum_{g \in G_A} \lambda(q) n_g g^y + \ldots$$

for some integer $\lambda(q) \not\equiv 0 \pmod q$. If G_A has only a single element g_0, then $m_g = \pm 1$, $n_g = \pm 1$, and the conclusion follows. Otherwise, let $g, h \in G_A$, $g \neq h$. From the above we have $\lambda(q)n_g \equiv m_g \pmod q$, $\lambda(q)n_h \equiv m_h \pmod q$ whence $m_h n_g \equiv m_g n_h \pmod q$. This holds for infinitely many primes $q \equiv y \pmod n$, whence $m_h n_g = m_g n_h$, and there exists λ such that $m_g = \lambda n_g$ for all $g \in G_A$. Let $\lambda = r/s$, $(r, s) = 1$ and suppose $|s| \neq 1$. Then some prime t divides s, and therefore t divides each n_g, contrary to the fact that the greatest common divisor of the $n_g = 1$. Hence $|s| = 1$. Now each m_h is divisible by r, so $|r| = 1$, and hence $\lambda = \pm 1$.

If G is cyclic, every automorphism is a power automorphism. We assume henceforth that G is a cyclic group of odd prime power order p^e. Then its automorphism group is cyclic and contains an automorphism mapping g into g^2.

Let

$$z(a) = G^{p^a} = \{g^{p^a} | g \in G\}.$$

The lattice of subgroups of G is a chain of characteristic subgroups:

$$G = Z(0) \supset Z(1) \supset \ldots \supset Z(e-1) \supset Z(e) = 1.$$

Define $C(a)$ to be the set difference $Z(a) - Z(a+1)$ for $e > a$, and $C(e) = Z(e) = 1$. We could alternatively define $C(a)$ as the set of g^t such that g is a generator of G and $t \equiv 0 \pmod{p^a}$, and $t \not\equiv 0 \pmod{p^{a+1}}$. We refer to the $C(a)$ as the *levels* of G.

Since the $Z(a)$ are characteristic subgroups of G, each $Z(a)$ and consequently each $C(a)$ is fixed under every automorphism of G.

If we define $G_A(a)$ as the intersection of G_A and $C(a)$, then by Theorem 1, if g_1 and g_2 are elements of $G_A(a)$ it follows that $m_{g_1} = \pm m_{g_2}$, for we can find an automorphism (y) which maps g_1 into g_2.

Let Y be the group of automorphisms of G. If A is a basis element of P, let Y_A be the subgroup of Y which leaves G_A fixed, that is

$$Y_A = \{(y) \in Y | A^{(y)} = \pm A\}.$$

We define the *spectrum of the set* G_A of the partition induced by P as $\mathrm{Sp}(G_A) = \{a | G_A(a) \neq 0\}$. Thus, the spectrum of a set is the collection of integers corresponding to levels intersected by the set. We define two basis elements B and D to be conjugate if $B = D^{(w)}$ for $(w) \in Y$.

If two basis elements are conjugate, their induced partition sets have the same spectra. Also, if two partition sets have intersecting spectra, there is an automorphism (y) of G mapping an element of one into an element of the other, and hence mapping the sets into each other. Thus, we can state:

LEMMA 1.1. *If the spectra of G_A and G_B intersect, then* $\mathrm{Sp}(G_A) = \mathrm{Sp}(G_B)$.

Now we will prove a corollary to Theorem 1.

COROLLARY. *Let G be a cyclic group of odd prime power order, and let P be a partition ring of G. There exists a basis for P such that if A is an element of this basis so is $A^{(y)}$ for any y prime to the order of G.*

Consider any basis for P. Choose $A_1, \dots, A_k$ a maximal set of elements of this basis such that no A_i is conjugate to $\pm A_j$ for $i \neq j$. Let B be the set of all distinct $A_i^{(y)}$ for $(y) \in Y$. Clearly, if no $A_i^{(w)} = -A_i^{(z)}$ for (w), $(z) \in Y$, then B is a basis for P. Suppose that $A_i^{(w)} = -A_i^{(z)}$ for (w), $(z) \in Y$. Then

$$A_i^{(wz^{-1})} = -A_i.$$

We will now show that this is impossible.

First, we introduce the following notation. If W is any expression of the form

$$\sum_{g \in G_W} m_g g$$

where G_W is a subset of the elements of G and the m_g are integers, we define $|W|$ as

$$\sum_{g \in G_W} m_g.$$

If K is a set of elements, we let $|K|$ be the cardinal of the set. For $0 < z < e$, we define

$$W(a) = \sum_{g \in G_W \cap C(a)} (m_g g).$$

Let basis element

$$A = \sum_{g \in G_A} m_g g.$$

By Theorem 1, there exist integers m_a, and integers $\alpha_g = \pm 1$ such that

$$A = \sum_{a \in Sp(G_A)} m_a \sum_{g \in G_A(a)} \alpha_g g.$$

Assume $A^{(u)} = -A$, $(u) \in Y_A$. First, we note that (u) and hence Y_A have even order, and second that for $g \in G_A$, precisely half of the α_g are -1.

Let b be the smallest integer in the spectrum of G_A. We may write $A = D + E$ where D is a linear combination of elements of $C(b)$ while E is a linear combination of elements of $Z(b+1)$. We note that $|D| = |E| = 0$.

Since $A^2(b)$ is a linear combination of conjugates of D, $|A^2(b)| = 0$. Thus

$$|(D+E)^2(b)| = |D^2(b)| + 2(|(DE)(b)|) + |E^2(b)| = 0.$$

Since $Z(b+1)$ is a group, $E^2(b) = 0$. Since no element of $C(b)$ is an element of $Z(b+1)$, every product gh for $g \in C(b)$, $h \in Z(b+1)$, is an element of $C(b)$ and $DE(b) = DE$, so $|(DE)(b)| = |DE| = 0$. Thus $|D^2(b)|$ must equal

zero. By computation, we will obtain a contradiction to this statement and hence show that $A^{(w)} = -A$ is impossible.

We have

$$D = A(b) = m_b \sum_{g \in G_A(b)} \alpha_g g.$$

Since Y_A acts transitively on $G_A(b)$, all the subgroups U_g leaving fixed an element $g \in G_A(b)$ have the same order u, and, for chosen g, each $g' \in G_A(b)$ appears as g^y for exactly u elements $(y) \in Y_A$. For each $(y) \in Y_A$, $A^{(y)} = \beta_y A$ where $\beta_y = \pm 1$. For $(y) \in U_g$, any g, comparison of the coefficients of g in A and $A^{(y)}$ shows that $\beta_y = +1$. Thus the β_y are equal for all those (y) carrying g into a given $g' = g^y$, and comparison of coefficients again shows that $\alpha_g' = \beta_y \alpha_g$. It follows that the term $\alpha_{g'} g'$ occurs exactly u times in the sum

$$\sum_{(y) \in Y_A} \beta_y \alpha_g g^y.$$

Hence, for any $g \in G_A(b)$, we have

$$D = \frac{m_b}{u} \sum_{(y) \in Y_A} \beta_y \alpha_g g^y.$$

Now we may write

$$\begin{aligned} D^2 &= \left(m_b \sum_{g \in G_A(b)} \alpha_g g\right) D \\ &= m_b \sum_{g \in G_A(b)} \left\{\alpha_g g \frac{m_b}{u} \sum_{(y) \in Y_A} \beta_y \alpha_g g^y\right\} \\ &= \frac{m_b^2}{u} \sum_{g \in G_A(b)} \sum_{(y) \in Y_A} \beta_y \alpha^2 g^{y+1} \\ &= \frac{m_b^2}{u} \sum_{(y) \in Y_A} \beta_y \left(\sum_{g \in G_A(b)} g^{y+1}\right). \end{aligned}$$

For $g \in C(b)$, we have $g^{y+1} \in C(b)$ if and only if $y + 1 \not\equiv 0 \pmod{p}$, and thus

$$|D^2(b)| = \frac{m_b^2}{u} |G_A(b)| \sum{}' \beta_y,$$

with summation over all $(y) \in Y_A$ such that $y \not\equiv -1 \pmod{p}$. To obtain a contradiction it will suffice to show that this sum is not zero.

The kernel of the natural map from the group Y, of order $p^{e-1}(p-1)$, onto the multiplicative group, of order $p-1$, of residues modulo p, has odd order p^{e-1}. Hence the intersection of this kernel with Y_A has odd order $p^{e'}$ dividing p^{e-1}. Since Y_A has even order, it contains elements mapping into -1, and hence exactly $p^{e'}$ of them. But $\Sigma' \beta_y$, as a sum of an odd number $|Y_A| - p^{e'}$ of terms $\beta_y = \pm 1$, cannot vanish.

We define the spectrum $S(B)$ of an element B of P as the sum of its distinct conjugates. We let $\bar{C}(a)$ be the sum of the elements of $C(a)$ and let $\bar{Z}(a)$ be the sum of the elements of $Z(a)$. We note that, since each $Z(a)$ is a group, for $b \geqslant a$, $\bar{Z}(a)\bar{Z}(b) = (|\bar{Z}(b)|)\bar{Z}(a)$, and since $\bar{C}(a) = \bar{Z}(a) - \bar{Z}(a+1)$ (where we define $Z(y)$ empty for $y > e$),

$$\bar{C}(a)\bar{C}(b) = (\bar{Z}(a) - \bar{Z}(a+1))(\bar{Z}(b) - \bar{Z}(b+1)).$$

If $b > a$, this is $(|\bar{C}(b)|)\bar{C}(a)$, while

$$\begin{aligned}\bar{C}(a)^2 = (|\bar{Z}(a)| - |(\bar{Z}(a+1)|)\bar{Z}(a) - (\bar{Z}(a) - \bar{Z}(a+1))(|\bar{Z}(a+1)|)\\ = (|\bar{C}(a)|)\bar{Z}(a) - (|\bar{Z}(a+1)|)\bar{C}(a).\end{aligned}$$

Thus, if S is an element of the form

$$\sum_{x=0}^{e} m_x\bar{C}(x)$$

then

$$\begin{aligned}S^2 &= \sum_{x=0}^{e}\left(m_x^2\bar{C}(x)^2 + 2\sum_{y=x+1}^{e} m_x m_y\bar{C}(x)\bar{C}(y)\right)\\ &= \sum_{x=0}^{e} m_x^2(|\bar{C}(x)|)\bar{Z}(x) - m_x^2(|\bar{Z}(x+1)|)\bar{C}(x) + 2\sum_{y=x+1}^{e} m_x m_y(|\bar{C}(y)|)\bar{C}(x)\\ &= \sum_{x=0}^{e} h_x\bar{C}(x)\end{aligned}$$

where

$$(1) \qquad h_x = \sum_{y=0}^{x} m_y^2(|\bar{C}(y)|) - m_x^2\sum_{y=x+1}^{e}|\bar{C}(y)| + 2\sum_{y=x+1}^{e} m_x m_y(|\bar{C}(y)|).$$

An immediate result of this computation is the following lemma.

LEMMA 1.2. *If*

$$S = \sum_{x=0}^{e} m_x\bar{C}(x)$$

then

$$S^2 = \sum_{x=0}^{e} h_x\bar{C}(x)$$

where, if $m_x = 0$, *then*

$$h_x = \sum_{y=0}^{x-1} m_y^2(|\bar{C}(y)|).$$

LEMMA 1.3. *If A is a basis element of P, then $S_p(G_A) = \{x|a < x < d\}$ for some $d \geqslant a$.*

Suppose not. Then there exists a smallest a such that there exists G_A with $a \in \mathrm{Sp}(G_A)$ and such that there exist d and c such that $d > c > a$ with

$d \in \mathrm{Sp}(G_A)$ and $c \notin \mathrm{Sp}(G_A)$. Then $c \in \mathrm{Sp}(G_B)$ for some basis element B, and by minimality of a, the minimal element of $\mathrm{Sp}(G_B)$ is greater than a. But, by previous calculations, if

$$S(B) = \sum_{x=0}^{e} m_x \bar{C}_x,$$

$$(S(B))^2 = \sum_{x=0}^{e} h_x \bar{C}(x)$$

where h_x is given by (1). Thus, since a is less than the minimal integer of $\mathrm{Sp}(G_B)$, $h_a = 0$. However, by Lemma 1.2,

$$h_a = \sum_{y=0}^{a-1} m_y^2 |\bar{C}(y)| > 0$$

since $m_c{}^2 > 0$. Hence the partition set of $(S(B))^2$ contains part but not all of the partition set of A, which contradicts the assumption that A is a basis element.

We have shown that the spectrum of a basis element is a set of consecutive integers. Now we will examine coefficients of the levels of the basis elements.

First, we note that if $S = S(A)$, S^2 and S have intersecting partition sets. For if b is the maximal integer in $\mathrm{Sp}(G_A)$,

$$S^2 = \sum_{x=0}^{e} h_x \bar{C}(x),$$

h_x given by (1), and $m_a = 0$ unless $a \in \mathrm{Sp}(G_A)$. Thus

$$h_b = \sum_{\substack{a \in \mathrm{Sp}G_A \\ a<b}} m_a^2 |\bar{C}(a)| + m_b^2 [|\bar{C}(b)| - |\bar{Z}(b+1)|].$$

If $e = b$, $h_b > 0$ since $\bar{Z}(b+1) = 0$, otherwise $|\bar{Z}(b)| = p|\bar{Z}(b+1)|$ and $\bar{C}(b) = \bar{Z}(b) - \bar{Z}(b+1)$, and hence

$$h_b = \sum_{\substack{a<b \\ a \in \mathrm{Sp}G_A}} m_a^2 |\bar{C}(a)| + m_b^2 (p-2) |\bar{Z}(b+1)| > 0.$$

If $\mathrm{Sp}(G_A) = a, a+1, \ldots d, d > a$ we may write

$$S(A) = S = \sum_{x=a}^{d+1} n_x \bar{Z}(x)$$

where $n_x = m_x - m_{x-1}$, $m_x = 0$ for $x < a$ or $x > d$, but since for $x < y$,

$$\bar{Z}(x)\bar{Z}(y) = (|\bar{Z}(y)|)\bar{Z}(x) = p^{e-y}\bar{Z}(x),$$

$$S^2 = \sum_{x=a}^{d} \left(n_x p^{e-x} + 2 \sum_{w=x+1}^{d+1} n_w p^{e-w} \right) n_x \bar{Z}(x) + n_{d+1}^2 p^{e-d-1} \bar{Z}(d+1).$$

Since the spectra of S and S^2 intersect, $S^2 = kS + \ldots$, where k is a nonzero integer. So for $x = a, a+1, \ldots, d$,

$$k n_x = \left(n_x p^{e-x} + 2 \sum_{w=x+1}^{d+1} n_w p^{e-w} \right) n_x.$$

But we know that $n_a = m_a$ is not equal to 0. If all the n_x, $a < x < d$ are zero, then the m_x are all equal and thus all $m_x = 1$. Suppose that not all m_x are equal. Let $H = \{h | a < h < d, n_h \neq 0\}$.

Let u and v, $u < v$ be two consecutive members of H. Then

$$kn_u = \left(n_u p^{e-u} + 2\sum_{w=u+1}^{d+1} n_w p^{e-w}\right) n_u,$$

$$kn_v = \left(n_v p^{e-v} + 2\sum_{w=v+1}^{d+1} n_w p^{e-w}\right) n_v.$$

Since u and v are consecutive in H, $n_u \neq 0$, $n_v \neq 0$ but $n_k = 0$ for $u < k < v$, and hence solving the above equations, we obtain $n_v = -p^{v-u} n_u$. Now $n_v = m_v - m_u$ since $m_{v-1} = m_u$. But $m_a \neq 0$ since $a \in \mathrm{Sp}(G_A)$, so that if there exists an integer larger than a in H we set $u = a$ and let v be the next smallest integer in H. Then $m_v = m_a(1 - p^{v-a})$.

Therefore, the sign of m_v is the negative of that of m_a and we can state:

LEMMA 1.4. *If*

$$S(A) = \sum_{x \in \mathrm{Sp}(G_A)} m_x \bar{C}(x)$$

and each m_x is positive, then all $m_x = 1$.

A basis element such that each m_v is positive will be called a *positive* basis element. If A is not a positive basis element, the above equations show that

$$H = \{a, h_1, h_2, \ldots, h_k\}$$

where $a < h_1 < \ldots < h_k$, $k > 0$, and

$$S = n_a \bar{Z}(a) + n_{h_1} \bar{Z}(h_1) + n_{h_2} \bar{Z}(h_2) + \ldots + n_{h_k} \bar{Z}(h_k) \tag{2}$$

where

$$n_{h_j} = (-1)^j (p^{h_j - a}) n_a$$

for $j = 1, 2, \ldots, k$.

A basis element with spectrum S as defined by (2) is called an *alternating* basis element.

LEMMA 1.5. *If A is an alternating basis element then*

$$S(A) = \sum_{x=0}^{d+1} n_x \bar{Z}(x),\ n_0 = \pm 1,$$

and if $0 < h_1, < \ldots < h_k$ are the elements of $\mathrm{Sp}(G_A)$ with $n_h \neq 0$, then

$$n_{h_j} = (-1)^j (p^{h_j}) n_0.$$

We must show that $n_0 \neq 0$. Suppose $n_0 = 0$. There exists a basis element $B \neq A$, such that G_B intersects $C(0)$, and therefore

$$S(B)^2 = \sum_{t \in \mathrm{Sp}(G_A)} m_t \bar{C}(t) + \ldots$$

where each $m_t > 0$ by Lemma 1.2. Then $S(B)^2 = kA + \dots$, and the coefficients in A must have the same sign as k, so A must be a positive element. This shows that $n_0 \neq 0$ and thus $a = 0$ in (2) and n_0 divides all the n_h. Thus since the greatest common divisor of the m_g is 1 for g in the partition set corresponding to basis element A, n_0 must be $+1$ or -1. We will always choose $n_0 = 1$ for an alternating element of the canonical basis of its partition ring.

We now show that if A is a basis element then $A \neq S(A)$ implies $G_A \subseteq C(a)$ for some a.

LEMMA 2.1. *Let A be a basis element of a partition ring P of a group G of odd prime power order p^e. Let Y be the automorphism group of G, and Y_A the subgroup of Y leaving A fixed. If $[Y\colon Y_A]$ is not a power of p, there exists $(z) \in Y$ such that $G_A(a) \,.\, G_A^{(z)}(a) \subseteq C(a)$ for all a.*

Let (y) be a generator of Y, g a generator of G, and let a be an integer. $[Y\colon Y_A] = p^s b$ where $b | p - 1$ since Y has order $p^{e-1}(p-1)$ and $b \neq 1$ by hypothesis. Both Y and Y_A are cyclic so Y_A is generated by $y_A = y^{p^s b}$.

If $G_A(a)$ is empty, the result is trivial for any $(z) \in Y$.

If $G_A(a)$ is non-empty, $G_A(a)$ contains g^{vy^n} for some $(v) \in Y$. Then, for all $(z) \in Y$,

$$g^{vzy^n} \in G_A^{(z)}(a),$$

and $G_A(a)$ and $G_A^{(z)}(a)$ are closed under Y_A. Therefore the existence of a $z \not\equiv 0 \pmod p$ such that $G_A(a) \,.\, G_A^{(z)}(a) \subseteq C(a)$ is equivalent to the existence of a z such that for all $m, n, p \nmid (vy_A^m + vzy_A^n)$, hence is equivalent to the existence of a z such that for all $r, p \nmid (y_A^r + z)$ or $y_A^r \not\equiv -z \pmod p$. This condition is clearly independent of a.

But

$$y_A = y^{p^s b} \equiv y^b \pmod p$$

and y is primitive of order $p - 1 \pmod p$ while y_A is of order $(p-1)/b \pmod p$ and hence has order less than $p - 1$. Thus we may choose $-z$ from the residues which do not appear as powers of y_A.

Next, we mention a well-known lemma.

LEMMA 2.2.

$$\text{If} \quad a^{p^i} \equiv b^{p^i} \pmod p$$

for some $i > 0$, p a prime, then

$$a^{p^n} \equiv b^{p^n} \pmod{p^{n+1}}$$

for all non-negative integers n. (*See* **2**, Theorem 4.5, vol. 1.)

LEMMA 2.3. *Under the hypotheses of Lemma* 2.1, if

$$[Y\colon Y_A] = p_s,\ s > 0, \text{ then } G_A(a)^2 \cap C(a+1) = 0$$

for all a.

Let $g^{p^s u}$ be an element of $G_A(a)$. Then an element of $G_A(a)^2$ would be of the form

$$g^{p^s}(u)\,(y_A^{n_1} + y_A^{n_2}).$$

If this element is contained in $C(a+1)$, $p\|(y_A{}^{n_1} + y_A{}^{n_2})$ (where we use $p^k\|u$ to mean that p^k is the highest power of p dividing u).

Let $p|y_A{}^{n_1} + y_A{}^{n_2}$. Then $y_A{}^{n_1} \equiv -y_A{}^{n_2} \pmod p$. But

$$y_A = y^{p^s}$$

(where y generates Y) so

$$y^{p^s n_1} \equiv -y^{p^s n_2} \pmod p.$$

Now Lemma 2.2 shows that

$$y^{p^s n_1} \equiv -y^{p^s n_2} \pmod{p^{s+1}}$$

and since $s > 0$, $p^2|(y_A{}^{n_1} + y_A{}^{n_2})$, which means that the element does not lie in $C(a+1)$.

Lemma 2.4. *If A is a basis element of a partition ring of a cyclic group G of odd prime power order and G_A meets more than one level of G, then G_A is a union of levels.*

To show this, we need only show that $Y_A = Y$. Suppose $Y_A \neq Y$; since by Lemma 3.1 the spectrum of A is consecutive and G_A meets more than one level, there exists a such that $G_A(a) \neq 0$, $G_A(a+1) \neq 0$. Let $[Y: Y_A] = p^s b$, $(b, p) = 1$. First, assume $b \neq 1$. Choose z by Lemma 2.1 such that $G_A(h)G_{A^{(z)}}(h) \subseteq C(h)$ for all h. Let $B = A^{(z)}$.

$$(AB)(a) = A(a)B(a) + A(a)[B(a+1) + \ldots + B(d)] + B(a)[A(a+1) + \ldots + A(d)].$$

But since $|A(h)| = |B(h)|$, by Theorem 1,

$$\frac{|AB(a)|}{|A(a)|} = |A(a)| + 2[|A(a+1)| + \ldots + |A(d)|]$$

while similarly

$$\frac{|AB(a+1)|}{|A(a+1)|} = |A(a+1)| + 2[|A(a+2)| + \ldots + |A(d)|].$$

We chose z so that $G_A(a)\,G_B(a) \subseteq C(a)$, thus the spectra of AB and A intersect, and hence

$$AB = \sum_w k_w A^{(w)}$$

with distinct values of k_W, and the sum of the coefficients in AB of the elements in a given level is in a fixed proportion to those in A. Thus the left sides of the two equations above must be equal and, by subtracting, we obtain $|A(a)| + |A(a+1)| = 0$. But, by Theorem 1,

$$\frac{|G_A(a)|}{|G_A(a+1)|} = \frac{|C(a)|}{|C(a+1)|} = p$$

(where for any set S we shall write $|S| = |\bar{S}|$) and since $|A(a)| \neq 0$, we have $m_{a+1} = -pm_a$.

We have shown, however, that if A is a positive element $m_{a+1} = m_a$, and if A is an alternating element $m_{a+1} = -(p-1)m_a$, so we have obtained a contradiction.

By the above reasoning $b = 1$, whence the hypothesis that $Y_A \neq Y$ implies that $s > 0$. Now, by Lemma 2.3, for all h, $G_A(h)^2 \cap C(h+1) = 0$.

Let $g \in G_A(h)$. Then, an element of $G_A(h)^2$ is of the form $g^{(1+y_A^k)}$ and this is an element of $C(h)$ unless $p|1 + y_A{}^k$, that is, unless $y_A{}^k \equiv -1$ modulo p.

Let W be the set of elements of Y_A which are congruent to -1 modulo p. W is non-empty since Y_A is of even order (divisible by $p-1$) and hence is equal in order to the subgroup of Y_A of elements congruent to 1 modulo p. But this is a subgroup of index $p-1$ in Y_A, and hence

$$|W| = \frac{1}{p-1}(|Y_A|).$$

Thus

$$|G_A^2(h) \cap C(h)| = \frac{p-2}{p-1}|G_A(h)^2|.$$

So

$$\frac{|A^2(a)|}{|A(a)|} = \frac{p-2}{p-1}[|A(a)|] + 2[|A(a+1)| + \ldots + |A(d)|]$$

while

$$\frac{|A^2(a+1)|}{|A(a+1)|} = \frac{p-2}{p-1}[|A(a+1)|] + 2[|A(a+2)| + \ldots + |A(d)|]$$

and

$$\frac{p-2}{p-1}|A(a)| + \frac{p}{p-1}|A(a+1)| = 0 \quad \text{or} \quad m_{a+1} = -(p-2)m_a,$$

which also contradicts previous lemmas. Thus the lemma is proved.

Thus we have proved:

THEOREM 2. *Let A be a basis element of a partition ring of a cyclic group of odd prime power order. If $Y_A \neq Y$, then $G_A \subset C(a)$ for some a, and all m_g are* 1. *If $Y_A = Y$ then G_A is a union of consecutive levels. If all m_g are positive, then all m_g are* 1. *If not all m_g are positive, then $C(0) \subset G_A$ and A is an alternating basis element.*

Next, we examine some relations between sets which intersect consecutive levels.

Let G be a cyclic group of odd prime power order p^e and let Y be the automorphism group of G. A non-empty subset J of G is called a basic set if it is the set of all images of an element g of G under a subgroup of Y. The largest subgroup Z of Y such that $J = \{g^y | y \in Z\}$ is called the automorphism group of J.

We will now state three lemmas concerning basic sets and the sums of their elements. Let G_A be a basic set contained in $C(a)$, $0 < a < e - 1$ and let Y_A be the automorphism group of G_A. Let $[Y: Y_A] = p^s b$ where $s > 0$ and $b | p - 1$.

LEMMA 3.1. *G_A is a union of $d = (p-1)/b$ cosets of G^{p^a} modulo $G^{p^{a+s+1}}$.*

Let H be a coset contained in G_A. We define $H^{[p]}$ to be the coset containing the pth powers of the elements of H. Let A be the sum of all elements of G_A and $A^{[p]}$ the sum of all elements of the cosets $H^{[p]}$ for cosets $H \subseteq G_A$.

LEMMA 3.2.

$$A^p \equiv |G^{p^{a+s+1}}|^{p-1} A^{[p]} \pmod{p |G^{p^{a+s+1}}|^{p-1}}.$$

Let G_B be a basic subset contained in

$$\bigcup_{H \subseteq G_A} H^{[p]}$$

and Y_B be the automorphism group of G_B. Let B be the sum of all elements of G_B.

LEMMA 3.3. *AB is a sum of conjugates, under Y, of A if and only if $G_B \cap H^{[p]} \neq 0$ for each coset $H \subseteq G_A$.*

Proof of Lemma 3.1. If (y) is a generator of Y, then $(Y_A) = (y^{p^s b})$ is a generator for Y_A. As a generator of Y, (y) is transitive on all levels of G and hence on $C(a)$. The order of $C(a)$ is $p^{e-a-1}(p-1)$, and thus this is the order of (y) on $C(a)$. Let

$$(z) = (y_A)^d = (y^{p^s(p-1)}).$$

Then (z) has order

$$p^{e-a-s-1} = p^{e-a-1}(p-1)/p^s(p-1)$$

on $C(a)$. By Lemma 2.2, $z \equiv 1 \pmod{p^{s+1}}$, and hence, for any $h = g^{p^a}$ in G^{p^a}, $h^{(z)} \equiv h \pmod{G^{p^{a+s+1}}}$. Since the order of (z) on $C(a)$ is equal to the number of elements in a coset modulo $G^{p^{a+s+1}}$ and (z) maps each coset into itself, (z) is transitive on each coset contained in $C(a)$. It follows that G_A is a union of cosets.

If $h = g^{p^a} \in G_A$, then

$$h^{p y_A^i} \equiv h^p \pmod{G^{p^{a+s+1}}}$$

$$p y_A^i \equiv p \pmod{p^{s+1}}$$

and since $s > 0$, $y_A{}^i \equiv 1 \pmod{p}$. But

$$y_A = y^{p^s b} \equiv y^b \pmod{p},$$

whence $d|i$. Thus, for $h \in G_A$ the elements

$$h, h^{\nu_A}, \ldots, h^{\nu_A^{d-1}}$$

lie in distinct cosets, say $H_0, H_1, \ldots, H_{d-1}$, and further, the cosets

$$H_0^{[p]}, \ldots, H_{d-1}^{[p]}$$

are distinct. This proves that $G_A = H_0 \cup H_1 \cup \ldots \cup H_{d-1}$, a union of d cosets.

Proof of Lemma 3.2. Let

$$S_i = \sum_{g \in H_i} g, \quad S_i^{[p]} = \sum_{g \in H_i[p]} g$$

whence

$$A = \sum_{i=0}^{d-1} S_i, \quad A^{[p]} = \sum_{i=0}^{d-1} S_i^{[p]}.$$

Now

$$A^p = \left(\sum S_i\right)^p = \sum \binom{p}{k_0, \ldots k_{d-1}} S_0{}^{k_0} S_1{}^{k_1} \ldots S_{d-1}^{k_{d-1}},$$

and

$$S_0{}^{k_0} \ldots S_{d-1}^{k_{d-1}} = |G^{p^{a+s+1}}|^{p-1} S$$

for S the sum of the elements of some coset H. Hence,

$$\text{modulo } p|G^{p^{a+s+1}}|^{p-1},$$

$$A^p \equiv \sum_{i=0}^{d-1} S_i^p \equiv \sum_{i=0}^{d-1} |G^{p^{a+s+1}}|^{p-1} S_i{}^{[p]} = |G^{p^{a+s+1}}|^{p-1} A^{[p]}.$$

Proof of Lemma 3.3. For $i = 0, 1, \ldots, d-1$, let $|G_B \cap H_i{}^{[p]}| = m_i$. If $m_i, m_j \neq 0$, there is an automorphism in Y_B mapping $G_B \cap H_i{}^{[p]}$ onto $G_B \cap H_j{}^{[p]}$, whence $m_i = m_j$. Thus, if some non-zero $m_i = m$, each $m_j = 0$ or m, for $j = 0, 1, \ldots, d-1$.

For $h \in H_0$, $H_i H_j{}^{[p]}$ contains

$$h^{\nu_A^i + p\nu_A^j} = h^{\nu_A^i(1+p\nu_A^{j-i})},$$

hence

$$H_i H_j^{[p]} = H_i{}^{(1+p\nu_A^{j-i})},$$

a conjugate of H_i. Moreover, the

$$H_i{}^{(1+p\nu_A^k)}$$

for $0 < i, k < d - 1$ are distinct, for

$$h^{y_A^i(1+py_A^k)} \equiv h^{y_A^{i'}(1+py_A^{k'})} \pmod{G^{p^{a+s+1}}}$$

implies

$$y_A^i(1 + py_A^k) \equiv y_A^{i'}(1 + py_A^{k'}) \pmod{p^{s+1}},$$

hence, since

$$s + 1 > 2, y_A^i \equiv y_A^{i'} \text{ and } y_A^k \equiv y_A^{k'}$$

modulo p, and thus $i = i'$, $k = k'$.

Thus

$$AB = \sum_{i=0}^{d-1} S_i \sum_{j=0}^{d-1} \left(\sum_{g \in H_j^{[p]} \cap G_B} g \right) = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} m_j S_i^{(1+py_A^{j-1})},$$

where $m_i = |H_i^{[p]} \cap G_B|$. Since

$$A^{(1+py_A^k)} = \sum_i S_i^{(1+py_A^k)},$$

AB is a sum of such conjugates if and only if all $m_j = m \neq 0$, that is, since $G_B \neq 0$, if G_B meets each $H_i^{[p]}$.

A result of these lemmas is the following theorem.

THEOREM 3. *Let G be a cyclic group of odd prime power order p^e. Let A be a basis element of a partition ring P of G such that the number of distinct conjugates of A is greater than or equal to p. Then G_A is entirely contained in some level $C(a)$ of G, and, if G_B is a partition set intersecting $C(a + 1)$, G_B is contained in $C(a + 1)$ and $|G_A| = p^t|G_B|$ for t an integer greater than or equal to zero.*

Proof of Theorem 3. That A has more than $p - 1$ conjugates implies that the index $[Y: Y_A]$ of the automorphism group of A is greater than p, hence $[Y: Y_A] = p^s b$ for some $s > 0$ and $b | p - 1$. By Theorem 2, $G_A \subseteq C(a)$ for some a, and since $C(e - 1)$ has only $p - 1$ elements, $0 < a < e - 1$. Writing $G_A = H_0 \cup \ldots \cup H_{d-1}$ in accordance with Lemma 3.1 we note that, using Lemma 3.2, we can show that $A^{[p]}$ is a linear combination of basis elements by an argument identical with that used in the first part of the proof of Theorem 1. But the partition set of $A^{[p]}$ is contained in $C(a + 1)$, and hence $C(a + 1)$ is a sum of partition sets of basis elements. Since all basis elements contained in a level are conjugate, and

$$H_0^{[p]} \cup \ldots \cup H_{d-1}^{[p]}$$

is a union of partition sets of basis elements, we may assume G_B is a set of this union. By Lemma 3.3, since G_B is a basic subset of G, $|G_B| = md$ where $m | (|H_i^{[p]}|)$. Since $|H_i^{[p]}| = p^{e-a-s-1}$ and $|G_A| = p^{e-a-s-1}d$, $|G_A| = p^t|G_B|$ for some $t > 0$.

We now show that the necessary conditions for a set Σ of elements

$$A = \sum_{g \epsilon G_A} m_g g$$

of the group ring of a cyclic group G of odd prime power order p^e, where the G_A constitute a partition π of G to be a canonical basis for a reduced integral partition ring of the group, as stated in Theorems 1, 2, and 3, are also sufficient. These conditions are as follows.

The partition π consists of sets of two sorts:

(*i*) $G_A = C(a) \cup C(a+1) \cup \ldots \cup C(d) = G^{p^a} - G^{p^{d+1}}, d > a;$

(*ii*) $G_A \subseteq C(a) = G^{p^a} - G^{p^{a+1}}.$

The sets of type (ii) are subject to the two conditions:

(iii) $G_A \subseteq C(a)$ implies that $G_A{}^{(y)} \in \pi$ for all $(y) \in Y$;

(iv) If $C(a)$ is a union of k sets G_A of π and $k \geqslant p$;

then $C(a+1)$ is a union of sets G_B of π, and $|G_A| = p^t|G_B|$ for t a non-negative integer.

(v) The elements of Σ are of two sorts:

$$A = \sum_{g \epsilon G_A} g,$$

(we will call such an element positive);

(b) Possibly, for a single

$$G_A = C(0) \cup C(1) \cup \ldots \cup C(d), d > 0,$$

A is an alternating element as defined earlier. The sufficiency of these conditions is asserted by the following theorem.

THEOREM 4. *Let G be a cyclic group of odd prime power order p^e. If π is a partition of G and Σ a set of elements of G such that π and Σ satisfy conditions* (i)–(v) *above, then Σ is a canonical basis for a partition ring of G.*

We need only show that if A and B are elements of the given set, Σ, then AB is a sum of elements of Σ. We may assume that the least element of the spectrum of G_A is less than or equal to the least element of the spectrum of G_B. Let Y_A be the automorphism group of G_A.

Case I. Let $Y = Y_A$ and let $\mathrm{Sp}(G_A) = \{a, \ldots, d\}$, $d \geqslant a$. Suppose $B = A$. Then $A^2 = kA + n\hat{Z}(d+1)$ for some k and n, and $\hat{Z}(d+1)$ is clearly a sum of elements of Σ such that all $m_g = 1$. If $B \neq A$ then $G_B \subseteq Z(d+1)$ and $BA = |B|A$.

We have considered the case in which $Y_A = Y$. Now we may assume that Y_A is a proper subgroup of Y, and thus $G_A \subset C(a)$ for some a.

Case II. Let $1 < [Y: Y_A] < p$. We consider two subcases:

Subcase II.1. Let G_B not intersect $C(a)$. Then since, by the construction employed in the proof of Lemma 3.1. G_A is a union of cosets of

$$G^{p^a} \text{ modulo } G^{p^{a+1}},$$

$G_B \subseteq G^{p^{a+1}}$ and hence $AB = |B|A$.

Subcase II.2. Let G_B intersect $C(a)$. By (iii) $B = A^{(v)}$ for some $(v) \in Y$. G_A must be a union of cosets modulo $G^{p^{a+1}}$ since $[Y: Y_A] < p$. If $G_A = H_0 \cup H_1 \cup \ldots \cup H_{d-1}$ then

$$G_B = H_0^{(v)} \cup H_1^{(v)} \cup \ldots \cup H_{d-1}^{(v)}.$$

Let

$$S_i = \sum_{g \in H_i} g.$$

Then

$$AB = A\,A^{(v)} = \sum_{i=0}^{d-1} S_i \sum_{j=0}^{d-1} S_j^{(v)} = |S_j| \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} S_i^{(1+vy_A^{j-i})}$$

$$= |S_j| \sum_{k=0}^{d-1} \sum_{i=0}^{d-1} S_i^{(1+vy_A^k)}.$$

If $vy_A^k \equiv -1 \pmod p$ then $S^{(1+vy_A^k)}$ is the sum of the elements of the unit coset $G^{p^{a+1}}$ and hence a sum of elements of Σ. If $vy_A^k \not\equiv -1 \pmod p$ then

$$\sum_{i=0}^{d-1} S_i^{(1+vy_A^k)}$$

is a conjugate of A. Thus AB is a sum of elements of Σ.

Case III. The index $[Y: Y_A] > p$. Then $[Y: Y_A] = p^s b$ where $s > 0$ and $b|p-1$. Let $d = (p-1)/b$. By Lemma 3.1, G_A is a sum of cosets of

$$G^{p^a} \text{ modulo } G^{p^{a+s+1}}.$$

We will retain the convention that $G_A = H_0 \cup H_1 \cup \ldots \cup H_{d-1}$, where

$$H_0^{(y_A^i)} = H_i.$$

We define $H^{(t)}$ to be the coset containing the t^{th} powers of the elements of H and $S^{(t)}$ to be the sum of the elements of $H^{(t)}$. Let $h \in H_0$ and let H_i and H_j be distinct cosets contained in G_A. If $H_i^{(t)}$ and $H_j^{(t)}$ are not distinct then

$$h^{ty_A^i} \text{ and } h^{ty_A^j}$$

are elements of the same coset and $ty_A^i \equiv ty_A^j \pmod{p^{s+1}}$. This implies that $p^{s+1}|t$, for otherwise we obtain $y_A^i \equiv y_A^j \pmod p$ contradicting the assumption of distinctness of cosets. We can now define $A^{(t)}$ as

$$\sum_{i=0}^{d-1} S_i^{(t)},$$

and note that if $t \equiv 0 \ (p^{s+1})$,

$$A^{(t)} = d\Big(\sum_{g\in G^{p^{a+s+1}}} g\Big),$$

while otherwise $A^{(t)}$ is a sum of distinct coset sums.

We will now prove two lemmas, under the assumptions of Case III.

LEMMA 4.1. *For $0 < k < s$, $C(a + k)$ is a union of at least p sets of π, each of which meets precisely d cosets of G^{p^a} modulo $G^{p^{a+s+1}}$ and these cosets are conjugate under Y_A.*

LEMMA 4.2. *The sum of $A^{(t)}$ is a sum of elements of Σ.*

Proof. of Lemma 4.1. We know that $|C(a)| = p^{e-a-1}(p-1)$ and

$$|G_A| = \frac{|C(a)|}{[Y:Y_A]} = p^{e-a-s-1}\,d.$$

By (iv), if G_{Ak} is a set of π contained in $C(a+k)$ and

$$\frac{|C(a+k)|}{|G_{Ak}|} > p,$$

then $C(a+k+1)$ is a union of sets of π and $|G_{Ak+1}| < |G_{Ak}|$, and since $|C(a+k)| = p|C(a+k+1)|$ we obtain

$$\frac{|C(a+k+1)|}{|G_{Ak+1}|} > \frac{|C(a+k+1)|}{|G_{Ak}|} = \frac{|C(a+k)|}{p|G_{Ak}|} > p^{s-k-1} \text{ for } 0 < k < s.$$

But since $|G_{Ak}| = p^t|G_{Ak+1}|$ for $t > 0$, and

$$[Y\colon Y_{Ak}] = \frac{|C(a+k)|}{|G_{Ak}|},$$

we may write $[Y\colon Y_{Ak}] = p^{s-k+\epsilon}b$ for $\epsilon > 0$.

If (y) is a generator for Y then $(y_A) = (y^{p^s b})$ is a generator for Y_A and $(y_{Ak}) = (y^{p^{s-k+\epsilon}b})$ is a generator for Y_{Ak}. Let H be a coset contained in G_A and $h \in H$. Now $h^{p^k} \in H^{(p^k)}$, and we examine the effect of (y_A) and (y_{Ak}) on this coset.

If $\epsilon > k$, from $y^{p^{\epsilon-k}} \equiv y \pmod p$, by Lemma 2.2, we obtain

$$y^{p^{s-k+\epsilon}} \equiv y^{p^s} \pmod{p^{s+1}}.$$

If $k > \epsilon$, from $y^{p^{k-\epsilon}} \equiv y \pmod p$ we obtain

$$y^{p^s} \equiv y^{p^{s-k+\epsilon}} \pmod{p^{s-k+\epsilon+1}}.$$

In either case

$$y^{p^s} \equiv y^{p^{s-k+\epsilon}} \pmod{p^{s-k+1}},$$

and thus

$$p^k y^{p^s b} \equiv p^k y^{p^{s-k+\epsilon}b} \pmod{p^{s-1}}.$$

Hence (y_A) and (y_{Ak}) map an element of $H^{(p^k)}$ into the same coset and thus they permute the cosets contained in $C(a+k)$ in the same way. It follows

that G_{A_k} meets precisely those cosets that belong to some family of d cosets conjugate under Y_A. This completes the proof of Lemma 4.1.

Proof of Lemma 4.2. By (v), since $\bar{C}(a)$ is not contained in an alternating element, each element of Σ with partition set contained in G^{p^a} is the sum of the elements of its partition set. Since, by Lemma 4.1, $C(a + s)$ is a union of elements of π, by (i) and (ii), $G^{p^{a+s+1}}$ is a union of elements of π, and hence for $t \equiv 0 \pmod{p^{s+1}}$, $A^{(t)}$ is a sum of elements of Σ.

If $p^{s+1} \nmid t$, $H_0^{(t)}$ is contained in some $C(a + k)$, $0 \leqslant k \leqslant s$, and by Lemma 4.1,

$$H_0^{(t)} \cup \ldots \cup H_{d-1}^{(t)}$$

is a union of sets of π. Hence $A^{(t)}$ is a sum of elements of Σ. This completes the proof of Lemma 4.2.

Let $G_B \in \pi$ be the partition set of some $B \in \Sigma$ where b, the smallest element of the spectrum of G_B, satisfies $b \geqslant a$. Then

$$B = \sum_{g \in G_B} g.$$

Subcase III.1. $b = a$. Then by (iii), $G_B = G_A^{(z)}$ for some $(z) \in Y$, whence

$$B = \sum_{j=0}^{d-1} S_j^{(z)}.$$

Now $|G_B \cap H_i^{(z)}| = |G^{p^{a+s+1}}|$ where $0 \leqslant i < d$ and

$$AB = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} S_i S_j^{(z)} = |G^{p^{a+s+1}}| \sum_{j=0}^{d-1} \sum_{i=0}^{d-1} S_i^{(1+zv_A^{j-1})}$$

$$= |G^{p^{a+s+1}}| \sum_{j=0}^{d-1} A^{(1+zv_A^j)}$$

which, by Lemma 4.2, is a sum of elements of Σ.

Subcase III.2. $a < b \leqslant a + s$. Then $C(b)$ is a union of conjugates of G_B, and writing $b = a + k$, we know that G_B meets $H_0^{(t)}$ for some t where $p^k \| t$, whence, by Lemma 4.1, G_B meets precisely the d cosets

$$H_0^{(t)}, H_1^{(t)}, \ldots, H_{d-1}^{(t)}.$$

By an earlier argument, see proof of Lemma 3.3, all $|H_i^{(t)} \cup G_B| = m$ for a fixed $m > 0$. Thus

$$AB = \sum_{i=0}^{d-1} S_i \Bigg(\sum_{j=0}^{d-1} \sum_{g \in H_j^{(t)}} g \Bigg) = m \sum_{j=0}^{d-1} A^{(1+tv_A^j)},$$

which again is a sum of elements of Σ.

Subcase III.3. If $b > a + s$, then B is an element of Σ with partition set contained in $G^{p^{a+s+1}}$. Then $BA = |B|A$ and the theorem is proved.

In view of Theorem 4, in order to list all reduced integral partition rings of a cyclic group G of odd prime power order p^e, it suffices to list all proper

partitions of G, that is, partitions satisfying (i)–(iv), and, in the case of partitions in which $C(0)$ is properly contained in a partition set, to list the possible ways in which an alternating element may occur.

If π is a proper partition of G, the restriction π' of π to G^p is a proper partition of the subgroup G^p. Thus, a proper partition π of G can be obtained from a proper partition π' of G^p in at most two ways:

1. $C(0)$ may be partitioned and the sets of this partition, together with those of π' will form π.

2. If $C(1)$, the lowest level of G^p, is contained in a set of π', $C(0)$ may be adjoined to this set to extend π' to π.

A partition formed by the first procedure must be made in such a manner that conditions (i)–(iv) are satisfied. This can be done as follows.

Let G_A be a set of the partition of $C(a)$. By condition (iii) G_A must be a basic set and from condition (iv) it follows that any such G_A for which

$$\frac{|C(0)|}{|G_A|} < p$$

yields a proper partition, while such a partition with

$$\frac{|C(0)|}{|G_A|} \geqslant p$$

is permissible just in case $C(1)$ contains some G_B and $|G_A| = p^s|G_B|$ for some $s \geqslant 0$. Any partition ring formed in this manner must contain only positive elements by (v) and hence is fully determined.

If the partition π is formed in the second manner, π contains a set $G_A = C(0) \cup C(1) \cup \ldots \cup C(d)$ for some $d > 0$. Then by (v), the element A of the partition ring with partition set G_A can be taken as a positive element or an alternating element. By the definition of an alternating element, the coefficient $m(0)$ of the elements of $C(0)$ must equal one, while the coefficients $m(a)$ of levels $C(a)$, $0 < a \leqslant d$ may each be chosen positive or negative, and by Lemma 1.5 the signs of these coefficients determine their values.

Since the group containing one element has only one partition ring, we have established an inductive procedure for finding all partition rings of cyclic groups of odd prime power order p^e.

References

1. B. Jónsson, and A. Tarski, *Representation problems for relation algebras*, Bull. Amer. Math. Soc. *54*, (1948), 80.
2. W. J. Leveque, *Topics in number theory*, 2 vols. (Reading, Mass.: Addison Wesley, 1956).
3. R. C. Lyndon, *The representation of relation algebras*, Ann. of Math. (2), *51* (1950), 707–729.
4. H. B. Mann, *On products of sets of group elements*, Can. J. Math., *4* (1952), 64–66.
5. A. Tarski, *On the calculus of relations*, J. Symb. Logic, *6* (1941), 73–89.
6. H. Zassenhaus, *The theory of groups* (New York: Chelsea, 1958).

University of Michigan and
Institute forDefence Analyses, Princeton

ON THE STRUCTURE OF SEMI-PRIME RINGS AND THEIR RINGS OF QUOTIENTS

JOACHIM LAMBEK

We are mainly interested in the study of prime and semi-prime rings and their rings of quotients. However, our argument proceeds largely in the category of modules (§ 1 to 4) and bimodules (§ 5 to 7).

After a brief description of the generalized rings of quotients introduced recently by Johnson, Utumi, and Findlay and the present author, we study a closure operation on the lattice of submodules of a module. For the lattice of left ideals of a ring, the concept of closed submodules reduces to the M-ideals of Utumi. The lattice of closed submodules of a module is always a complete modular lattice. We are specially interested in the case when it is a complemented lattice. This happens, in particular, when the singular submodule of Johnson and Wong vanishes. We consider the lattice of closed right ideals of a prime ring S and determine the maximal ring of right quotients of S in the case when this lattice has atoms. Our results for such prime rings are closely related to recent results by Goldie, Lesieur and Croisot, and Johnson.

All proofs in § 2 and § 3, concerning the closure operation on the lattice of all submodules of a module, have been carefully designed to carry over to an essentially different situation in § 5. There we study a closure operation, called *b-closure*, on the lattice of all submodules of a bimodule. This does not reduce to the original closure operation, even when the bimodule is converted into a right module. The connection between the two closure operations is rather exemplified by the following: Call a submodule *dense* (*b-dense*) if its closure (b-closure) is the whole module. Then an ideal in a ring is b-dense if and only if it is dense both as a right ideal and as a left ideal.

Each bimodule M possesses a *b-completion*, that is a largest bimodule in which M is b-dense. The b-completion of a ring S is also a ring and coincides with the so-called maximal ring of right and left quotients, first introduced in a special case by Utumi and defined in general by Johnson and Wong. The b-completion of a prime ring with non-zero socle is described symmetrically in terms of dual vector spaces.

The b-closed ideals of a semi-prime ring S are precisely its annihilator ideals. They form a complete Boolean algebra, which is isomorphic with the algebra of regular open sets in the prime ideal space of S. If S is also b-complete, the b-closed ideals are precisely the direct summands of S. This fact is exploited to obtain a structure theorem: Every such ring S is the direct sum of two

Received June 28, 1960.

rings C and C^*, where C is the complete direct product of b-complete prime rings and the lattice of annihilator ideals of C^* has no atoms.

The main results of § 5, § 6, and § 7 have been announced to the American Mathematical Society (Notices, 7 (1960), pp. 92 and 241).

I wish to thank Dr. Utumi for his careful reading and helpful criticism of the manuscript.

1. Survey of generalized rings of quotients.

1.1. If S is any associative ring, a right S-module M_S consists of an additive abelian group M and a mapping $(m, s) \to ms$ of $M \times S$ into M satisfying the obvious distributive and associative laws. Left modules are defined dually. The ring S gives rise, in an obvious way, to the right module S_S and the left module ${}_S S$.

A right module M_S is called *unitary* if S has a unity element 1 and $m1 = m$ for all $m \in M$. Every right module M_S can be converted into a unitary module $M_{S^\#}$ as follows: $S^\#$ is the ring consisting of the additive group $S \oplus Z$, Z the ring of integers, with multiplication defined by

$$(s + z)(s' + z') = (ss' + sz' + zs') + zz',$$

for $s, s' \in S$ and $z, z' \in Z$. One then puts

$$m(s + z) = ms + mz,$$

for $m \in M$, $s \in S$, and $z \in Z$.

1.2. Findlay and the present author **(5)** investigated a relation among three modules A_S, B_S, and C_S. They wrote $A < B(C_S)$ as an abbreviation for any of the following three equivalent statements:

(1) A_S is a submodule of B_S and, for any submodule E_S of $A_S - B_S$, $\mathrm{Hom}_S(E, C) = 0$. Here $A - B$ is the difference (or quotient) module of A modulo B.

(2) A_S is a submodule of B_S and, if $\phi \in \mathrm{Hom}_S(D, C)$, where D_S is any submodule of B_S and $A \subset \ker \phi$, the kernel of ϕ, then the image $\mathrm{im}\, \phi = 0$.

(3) A_S is a submodule of B_S and, for any $b \in B$ and any $0 \neq c \in C$, there exists an $s \in S$ and an integer z such that $bs + bz \in A$ and $cs + cz \neq 0$. If the modules in question are unitary, z can be taken to be 0.

1.3. If $A < B(B_S)$, B_S was called a *rational* extension of A_S. It was shown that any module M_S possesses a largest rational extension (*rational completion*) $\bar{M}_S$, unique up to isomorphism over M_S. $\bar{M}_S$ is *rationally complete* in the following sense: If $A < B(\bar{M}_S)$, then every $\phi \in \mathrm{Hom}_S(A, \bar{M})$ can be extended to a (unique) $\bar{\phi} \in \mathrm{Hom}_S(B, \bar{M})$. Two constructions of $\bar{M}_S$ were given:

(1) Let $M_S{}^I$ be the minimal injective extension **(4)** of M_S, then $\bar{M}_S$ consists of all those elements of $M_S{}^I$ which are annihilated by every endomorphism of $M_S{}^I$ which annihilates M_S.

(2) The right ideals D of $S^\#$ such that $D < S^\#(M_S)$ form a directed set under inclusion, and the additive groups $\mathrm{Hom}_S(D, M)$ form a direct system. Their direct limit is turned into an S-module $\bar{M}_S$ in a natural way. If M_S is unitary, one may replace the $S^\#$ of this construction by S.

1.4. Johnson and Wong **(19)** called a submodule L_S of M_S *large* if it has non-zero intersection with every non-zero submodule of M_S. They introduced the *singular* submodule $J(M_S)$ of a module M_S. It consists of all elements of M which annihilate a large right ideal of S. They showed that if $J(M_S) = 0$, then also $J(\bar{M}_S) = 0$ and $\bar{M}_S$ is injective. Moreover, the ring of endomorphisms of $\bar{M}_S$ is regular (in the sense of von Neumann) and injective as a right module.

1.5. If S is any associative ring, Q_S the rational completion of S_S, then Q is actually a ring extending S. Q coincides with the *maximal ring of right quotients* of S, previously defined by Johnson **(10)** and Utumi **(17)** in the following important cases.

Johnson's case. The singular submodule of S_S is actually an ideal, call it the *right singular ideal*. Johnson assumed that this ideal vanishes. He showed that the right singular ideal of Q then also vanishes and that Q is regular and injective as a right Q-module.

Utumi's case. Utumi assumed that, for any non-zero element s of S, $sS \neq 0$. It is, in fact, easily seen that this is a necessary and sufficient condition for Q to contain a unity element **(5**, 6.2**)**.

Among many other interesting applications, Utumi computed the maximal ring of left quotients of any primitive ring S with non-zero socle **(17**, 5.1**)**. Thus, let $V = eS$ be a minimal right ideal of such a ring, e an idempotent element of S **(9**, p. 57, Proposition 1**)**. Then $D = eSe$ is known to be a skew-field, and V is a vector space ${}_DV$. Utumi showed that $\mathrm{Hom}_D(V, V)$ is the maximal ring of left quotients of S.

1.6. For an integral domain S, the maximal ring Q of right quotients coincides with the classical field of quotients. If S is not an integral domain, there may also exist a "classical" ring of quotients. For example, if S is commutative, then this classical ring of quotients Q_{cl} consists of all ratios s/s', where $s \in S$ and s' is any *regular* element of S, in the sense that $s''s' \neq 0$ for any non-zero element s''. However, Q_{cl} may be smaller than Q. For instance **(2)**, if S is any Boolean ring, then $Q_{\mathrm{cl}} \cong S$, but Q is the Dedekind–MacNeille completion of S.

1.7. The ternary relation $A < B(C_S)$ has a number of properties, which are easily derived from the definition. We state them here for later reference.

P0. If $B - A \cong B' - A'$ and $C \cong C'$, then $A < B(C_S)$ implies $A' < B'(C_S')$.

P1. If $0 < C(C_S)$, then $C = 0$.

P2. If $A < B(C_S)$ and D_S is a submodule of C_S, then $A < B(D_S)$.

P3. If D_S is a submodule of B_S containing the submodule A_S (that is, $A \subset D \subset B$), then

$$A < B(C_S) \Leftrightarrow \text{both } A < D(C_S) \text{ and } D < B(C_S).$$

P4. $A < A(C_S)$.

P5. If $A < B(C_S)$ and $C < D(D_S)$, then $A < B(D_S)$.

Actually, it was shown in **(5)** that the second condition of P5 can be replaced by the weaker assumption that C_S is a large submodule of D_S. This stronger result will not be used here.

We mention also the following property, which has to do with change of rings **(5, 5.5)**.

(†) For any modules A_T, B_T, and C_T, if S is a subring of T such that $S < T(C_S)$, then

$$A < B(C_T) \Leftrightarrow A < B(C_S).$$

2. The lattice associated with a module. All modules are understood to be right S-modules.

2.1. Let A be a submodule of M. There is a largest submodule A^c of M containing A such that $A < A^c(M)$. This may be constructed as follows:

$$\begin{aligned} A^c &= \{m \in M \mid \ A < A + mS^\#(M_S)\} \\ &= \{m \in M \mid \ m^{-1}A < S^\#(M_S)\}. \end{aligned}$$

The second formula is due to Findlay. Here

$$m^{-1}A = \{x \in S^\# \mid \ mx \in A\}.$$

2.2. The assignment $c: A \to A^c$ is a closure operation on the lattice of all submodules of M. It has the following properties:

C1. $0^c = 0$.

C2. $(A \cap B)^c = A^c \cap B^c$.

C3. If $\phi \in \operatorname{Hom}_S(M, M)$, then $\phi(A^c) \subset (\phi A)^c$.

These correspond to A1, A2, and half of A3 of Johnson's "structures" on rings **(11)**.

Proof.

(C1) Since $0 < 0^c(M)$, therefore $0 < 0^c(0^c)$, by P2, hence $0^c = 0$, by P1.

(C2) Since c is a closure operation, $(A \cap B)^c \subset A^c \cap B^c$. To show the converse, observe that $A < A^c(M)$. From this we deduce that $A < A^c \cap (A + B)(M)$, by P3, that is $A < A + (A^c \cap B)(M)$, by the modular law. Now

$$(A + (A^c \cap B)) - A \cong (A^c \cap B) - (A \cap B),$$

by one of the isomorphism theorems of group theory. Therefore $A \cap B < A^c \cap B(M)$, by P0. Similarly we deduce from $B < B^c(M)$ that $A^c \cap B <$

$A^e \cap B^e(M)$. In view of P3, both these results together imply that $A \cap B < A^e \cap B^e(M)$.

(C3) Let $\phi \in \mathrm{Hom}_R(M, M)$, $K = A^e \cap \ker \phi$. Now $A < A^e(M)$, $\phi A^e \cong A^e - K$, $\phi A \cong (A + K) - K$, and $\phi A^e - \phi A \cong A^e - (A + K)$. By P3 and P0, $\phi A < \phi A^e(M)$, hence $\phi A^e \subset (\phi A)^e$, as required.

2.3. PROPOSITION. *The lattice $L(M)$ of closed submodules of M is a complete modular lattice, with set-intersection as meet.*

Proof. That we have a complete lattice follows from the fact that we have a closure operation. The join of two or more submodules of M is defined by

$$A \vee B = (A + B)^e, \quad \bigvee_{i \in I} A_i = \Big(\sum_{i \in I} A_i\Big)^e.$$

Finally, let A, B, and C be submodules of M and assume that $B \subset A$. Then

$$\begin{aligned} A \cap (B \vee C) &= A^e \cap (B + C)^e \\ &= (A \cap (B + C))^e \\ &= (B + (A \cap C))^e \\ &= B \vee (A \cap C), \end{aligned}$$

using C2 and the modular law for the lattice of all submodules of M.

2.4. A submodule K of M will be called *dense* if $K^e = M$. One easily verifies that every dense submodule is large.

LEMMA. *If K is dense in M, A any submodule of K, then $A^e \cap K$ is the closure of A in K.*

Proof. Since $A < A^e$, we have $A < A^e \cap K(K)$, by P2 and P3. Therefore $A^e \cap K$ is contained in the closure A^d of A in K.

Now $A < A^d(K)$, hence $A < A^d(M)$, by P5. Therefore $A^d \subset A^e$, and so $A^d \subset A^e \cap K$.

Note. We should really write $A^{e(M)}$ for A^e and $A^{e(K)}$ for A^d, but we have endeavoured not to make the notation too heavy.

2.5. The following partly generalizes a result by Utumi (**17**, Theorem 2).

PROPOSITION. *If K is dense in M, then $L(K)$ and $L(M)$ are isomorphic lattices under the inverse correspondences*

$$A \to A^e, \qquad B \to B \cap K,$$

where $A \in L(K)$ and $B \in L(M)$.

Proof. Again let d denote the closure operation in K. We observe that clearly $A^e \in L(M)$ and that $B \cap K \in L(K)$, since

$$(B \cap K)^d = B^d \cap K^d = B^e \cap K \cap K = B \cap K,$$

by C2 and the above lemma.

Next, we note that the two mappings are inverses. For $A^c \cap K = A^d = A$, by the lemma, and $(B \cap K)^c = B^c \cap K^c = B \cap M = B$, by C2 and the fact that K is dense in M.

Finally, we observe that the two mappings are meet-isomorphisms, hence lattice isomorphisms. For $B \cap B' \cap K = B \cap K \cap B' \cap K$ and $(A \cap A')^c = A^c \cap A'^c$.

2.6. PROPOSITION. *If K is a closed submodule of M, then any closed submodule of K is closed in M.*

Proof. Let A be a closed submodule of K, then $A < A^c(M)$, hence $A < A^c \cap K(K)$, by P2 and P3. Thus, $A^c \cap K \subset A$, which is closed in K. But $A \subset A^c$ and $A \subset K$, hence $A = A^c \cap K = A^c \cap K^c = (A \cap K)^c = A^c$, in view of C2.

2.7. Examples of closed submodules are the following submodules K of M:

(1) K is maximal such that $K \cap L = 0$, for some submodule L of M.

(2) $K = \{m \in M \mid Fm = 0\}$, for some subset F of $\mathrm{Hom}_S(M, M)$.

(3) K is a direct summand of M.

(4) K is rationally complete.

Indeed, (1) follows easily from the known fact that every dense submodule is large, (2) follows immediately from C3, and (3) is a special case of (2). Finally, assume that K is rationally complete, K^c its closure in M. Then K^c is a rational extension of K and therefore coincides with K, and so K is closed in M.

2.8. By the *socle* of a complete lattice we shall understand the join of all its *atoms*, that is its minimal non-zero elements.

PROPOSITION. *The socle of $L(M)$ is contained in every large closed submodule of M. It is mapped into itself by every endomorphism of M.*

Proof. Let A be an atom of $L(M)$, L a large closed submodule of M. Since $A \neq 0$, we have $A \cap L \neq 0$. Since A and L are closed, so is $A \cap L$. Since A is an atom, $A \cap L = A$, that is $A \subset L$. Thus L contains all atoms, hence their join.

Let $\phi \in \mathrm{Hom}_S(M, M)$ and let $\{A_i\}_{i \in I}$ be the set of all atoms of $L(M)$. By C3,

$$\phi\Big(\sum_{i \in I} A_i\Big)^c \subset \Big(\phi \sum_{i \in I} A_i\Big)^c \subset \Big(\sum_{i \in I} (\phi A_i)^c\Big)^c.$$

The result will follow if we show that the $(\phi A_i)^c$ are all 0 or atoms.

Let $A = A_i$ be any atom of $L(M)$. Any submodule of ϕA has the form ϕB, where $K \subset B \subset A$, K being the kernel of ϕ. Assume $B \neq 0$, then $\phi A - \phi B \cong A - B$. Now $B < A(M)$, hence $\phi B < \phi A(M)$, by P0. Therefore $\phi A \subset (\phi B)^c$, and so $(\phi A)^c \subset (\phi B)^c$.

Now let C be any closed submodule of M such that $0 \neq C \subset (\phi A)^c$. Since

ϕA is a large submodule of $(\phi A)^c$, $C \cap \phi A$ is a non-zero submodule of ϕA, hence has the form ϕB, where $B \neq 0$. By the above and C2,

$$(\phi A)^c \subset (C \cap \phi A)^c = C^c \cap (\phi A)^c = C.$$

Thus $(\phi A)^c$ is an atom, as remained to be shown.

2.9. PROPOSITION. *If M is any module, the socle of $L(M)$ is the closure of the discrete direct sum of some of its atoms. If $L(M)$ is a distributive lattice, then its socle is even the closure of the discrete direct sum of all the atoms.*

Proof. The argument for the first result is standard, for example, (**9**, p. 61). Indeed, let $\{A_i\}_{i\in I}$ be the set of all atoms of $L(M)$. By Zorn's lemma, one finds a maximal subset J of I such that, for all $i \in I$,

$$A_i \cap \bigvee_{j\in J-\{i\}} A_j = 0.$$

Now, for any $i \in I$, $A_i \cap \bigvee_{j\in J} A_j = 0$ or $= A_i$. By maximality of J, it is easily shown to be not 0, hence $A_i \subset \bigvee_{j\in J} A_j$. The first result now follows.

Next, assume that $L(M)$ is a distributive lattice. We will show that

$$A_i \cap \sum_{j\in I-\{i\}} A_j = 0,$$

for any $i \in I$. Thus, suppose that m belongs to the set denoted by the left side of this equation. Then there is a finite subset F of $I - \{i\}$ such that

$$m \in A_i \cap \sum_{j\in F} A_j \subset \bigvee_{j\in F} (A_i \cap A_j),$$

by the distributive law. Since $i \notin F$, A_i and A_j are distinct atoms, hence $A_i \cap A_j = 0$ for all $j \in F$. Therefore $m = 0$, as required.

3. Complemented lattices.

Unless otherwise stated, all modules are still assumed to be right S-modules.

3.1. Of special interest is the case where the lattice of closed submodules of a module is complemented.

LEMMA $L(M)$ *is complemented if and only if every large submodule of M is dense.*

We recall that a large submodule is one that has non-zero intersection with every non-zero submodule.

Proof. Assume $L(M)$ is complemented. This means that for every closed submodule A there is a closed submodule B such that $A \cap B = 0$ and $A \vee B = M$. Let L be any large submodule of M, then L^c will have a complement $K = K^c$. But then $K \cap L = 0$ and so $K = 0$. Hence $L^c = L^c \vee K = M$.

Conversely, assume the condition and let A be any closed submodule of M. Using Zorn's lemma, we find a maximal B such that $A \cap B = 0$. By 2.7(1),

B is closed. A well-known argument **(10)** now shows that $A+B$ is a large submodule of M. By assumption, $A + B$ is dense, hence its closure $A \vee B = M$. Thus B is a complement of A.

Our proof is now complete. Incidentally, we have shown:

If $L(M)$ is complemented and $A \in L(M)$, then any maximal submodule B of M such that $A \cap B = 0$ is a complement of A.

3.2. PROPOSITION. *If the lattice $L(M)$ associated with a module M is complemented then so is the corresponding lattice of any submodule and of any rational extension of M.*

Proof. Let $L(M)$ be complemented. If N is a rational extension of M, then $L(M) \cong L(N)$, by 2.5, hence $L(N)$ is also complemented.

Now let A be any submodule of M. Since A is dense in A^c, $L(A)$ and $L(A^c)$ are isomorphic, by 2.5. Thus it suffices to show that $L(K)$ is complemented, for any closed submodule K.

Let $B \in L(K) \subset L(M)$, by 2.6. Hence there exists $C \in L(M)$ such that $B \cap C = 0$ and $B \vee C = M$. We claim that $(C \cap K)^d$ is a complement of B in $L(K)$, where d is the closure operation for submodules of K.

Indeed, $C \cap K$ is a large submodule of $(C \cap K)^d$, hence $B \cap (C \cap K)^d = 0$. Moreover, by the modular law and C2,

$$(B + (C \cap K))^c = ((B + C) \cap K)^c = (B + C)^c \cap K^c = M \cap K = K.$$

Therefore, in view of 2.6,

$$\begin{aligned} K = (B + (C \cap K))^c &\subset ((B + (C \cap K))^d)^c \\ &= (B + (C \cap K))^d \subset (B + (C \cap K)^d)^d, \end{aligned}$$

hence the right side $= K$, as required.

3.3. THEOREM. *If M is rationally complete and $L(M)$ is complemented, then the following conclusions hold:*

(a) *Every closed submodule of M is a direct summand.*

(b) *For any submodule D of M, any $\phi \in \mathrm{Hom}_S(D, M)$ may be extended to an endomorphism of M.*

(c) *$F = \mathrm{Hom}_S(M, M)$ is a regular ring.*

(d) *The lattice $L(M)$ is isomorphic with the lattice of principal right ideals of F.*

(e) *F is injective as a right F-module.*

Proof.

(a) Let A be a closed submodule of M. By assumption, it has a complement B so that $A \cap B = 0$ and $A \vee B = M$. Consider the map $\phi \in \mathrm{Hom}_S(A + B, M)$ defined by $\phi(a + b) = a$. By rational completeness, this may be extended to $\psi \in \mathrm{Hom}_S(M, M)$. We have

$$\psi M = \psi(A + B)^c \subset (\psi(A + B))^c = A^c = A,$$

by C3. Thus, for any $m \in M$,

$$\psi^2 m = \psi(\psi m) = \phi(\psi m) = \psi m,$$

and so ψ is a decomposition operator.

(b) Let $D \subset M$, $\phi \in \mathrm{Hom}_S(D, M)$. By rational completeness of M, ϕ may be extended to $\phi' \in \mathrm{Hom}_S(D^c, M)$. By (a), D^c is a direct summand of M, hence ϕ' may be extended further to an element of $\mathrm{Hom}_S(M, M)$.

(c) Let $f \in F = \mathrm{Hom}_S(M, M)$. We observe that $K = \ker f$ is closed by 2.7(2). By (a), K is a direct summand, hence $M = K + H$ and $K \cap H = 0$. Thus f induces an isomorphism $g: H \to fH$. By (b), $g^{-1}: fH \to H$ may be extended to $f' \in F$. For any $k \in K$, $h \in H$, we thus have

$$ff'f(k + h) = ff'0 + fg^{-1}gh = fk + fh = f(k + h).$$

Therefore $ff'f = f$.

(d) This is proved like Johnson's theorem (**12**, II, 7.5), by showing that, for any idempotent $e \in F$, the principal right ideal eF of F determines the direct summand eM of M and vice versa. Thus $eM = (eF)M$ and $eF = \{f \in F | fM \subset eM\}$.

(e) This is proved like (**19**, Theorem 5).

3.4. Looking at the above proof, we find that the conditions of the theorem can be somewhat relaxed. Instead of rational completeness, it suffices to assume this:

For any submodule D of M, every $\phi \in \mathrm{Hom}_S(D, M)$ can be extended to some (necessarily unique) $\phi' \in \mathrm{Hom}_S(D^c, M)$.

It is easily seen that this condition is equivalent to the following:

M is mapped into itself by every endomorphism of the rational completion $\bar{M}$ of M.

3.5. *Examples.* The lattice $L(M)$ will be complemented if the singular submodule $J(M) = 0$. Johnson and Wong proved (c) and (e) for this important case. However this is not the only example.

The ring $S = Z_p$ of integers modulo the prime p may be regarded as a right Z-module. As such, its singular submodule $J(S_Z) = Z_p \neq 0$. Now, $L(S_Z)$ has only two elements, hence is trivially complemented. It can also be shown that S_Z is rationally complete.

Johnson and Wong (**19**, Theorem 5) have also shown that $\bar{M}$ is injective when $J(M) = 0$. This result cannot be generalized to the case when $L(M)$ is complemented. For Z_p, regarded as a Z-module, is not divisible.

3.6. We may ask when the lattice associated with a module consists of only two elements, that is, every non-zero submodule is dense. Goldie (**5**) has called a non-zero module *uniform* if every non-zero submodule is large. Thus, by 3.1, $L(M)$ has exactly two elements if and only if M is uniform and $L(M)$ is complemented.

Of special interest is the case when $S = Z$, the ring of integers.

PROPOSITION. *If M is an additive abelian group (Z-module), then $L(M)$ has exactly two elements if and only if M is cyclic of prime order or a subgroup of the additive group of rationals.*

We shall omit the proof, which depends on standard theorems in the theory of abelian groups.

3.7. A lattice is called *atomic* if every non-zero element contains ($\geqslant$) an atom, or minimal non-zero element.

PROPOSITION. *If the lattice $L(M)$ is complemented and atomic, then its socle is M. If the socle of $L(M)$ is M, then $L(M)$ is complemented.*

Proof. Assume that $L(M)$ is atomic and complemented. Let C be its socle, D a complement of C. Since $C \cap D = 0$, D contains no atoms, hence $D = 0$. Therefore $M = (C + D)^c = C^c = C$.

Conversely, suppose that $C = M$. By 2.8, every large, closed submodule of M coincides with M. By 3.1, $L(M)$ is complemented.

4. On prime rings.

4.1. An associative ring S is called *prime* if it has any one of the following equivalent properties:

(1) For any non-zero ideals A and B of S, $AB \neq 0$.
(2) For any non-zero elements s, s' of S, $sSs' \neq 0$.
(3) For any non-zero ideal A of S, $A^r = 0$.
(4) For any non-zero ideal B of S, $B^l = 0$. Here

$$A^r = \{s \in S | \ As = 0\}, \qquad B^l = \{s \in S | \ sB = 0\}$$

are the right and left annihilators of A and B respectively.

If S is a ring for which $S^l = 0$, it is well known (**5**, 6.4) and easily shown that an ideal A of S is dense as a submodule of S_S if and only if $A^l = 0$.

It follows that every two-sided ideal in a prime ring is dense.

LEMMA. *If S is a prime ring, the socle of $L(S_S)$ is either* 0 *or* S.

Proof. Suppose the socle of $L(S_S)$ is not 0. By 2.8, it is an ideal, hence dense. But, by definition, the socle is closed, hence it coincides with S.

4.2. If S is a prime ring, Q any ring of right quotients of S, then Q is also a prime ring. (It suffices to assume that S_S be a large submodule of Q_S.)

Indeed, let A and B be non-zero ideals of Q. Then $A \cap S$ and $B \cap S$ are non-zero ideals of S, hence $(A \cap S)(B \cap S) \neq 0$, and so $AB \neq 0$.

4.3. The following theorem owes its present form to a discussion with R. E. Johnson. (An independent proof of it was also found by Utumi.)

THEOREM. *If S is a prime ring such that the lattice $L(S_S)$ has non-zero socle, then its maximal ring of right quotients is a complete ring of linear transformations of a right vector space.*

Proof. We are given that S is a prime ring such that $L(S_S)$ has non-zero socle. Let Q be its maximal ring of right quotients, this is also prime, by 4.2. Moreover $L(Q_S) \cong L(S_S)$, by 2.5. We shall verify below that the closed submodules of Q_S are actually closed right ideals of Q, hence $L(Q_Q) = L(Q_S)$. Therefore, $L(Q_Q)$ also has non-zero socle, which must coincide with Q, by 4.1. Now, by 3.7, $L(Q)$ is complemented. Since $S^l = 0$, Q contains a unity element (see, for example (**5**, 6.2)). Therefore $Q \cong \mathrm{Hom}_Q(Q, Q)$, and this is a regular ring, by 3.3. Thus every principal right ideal of Q is a direct summand, hence a closed right ideal. Therefore, every atom of $L(Q_Q)$ is a minimal right ideal. Thus Q has non-zero socle. (The usual socle of Q is the socle of the lattice of all right ideals of Q.) Moreover Q_Q is rationally complete. By Utumi's theorem, mentioned in 1.5, $Q \cong \mathrm{Hom}_D(V', V')$, where D is a skew-field and V'_D is a right vector space.

4.4. The proof given above depended on the following lemma, which is implicit in the work of Utumi.

Lemma. *If Q is the maximal ring of right quotients of S then any closed submodule of Q_S is a closed right ideal of Q.*

Proof. Let A be a closed submodule of Q_S, and let $a \in A$. Take any $q' \in Q$ and $0 \neq q \in Q$. Since $S < Q(Q_S)$, we can find $x \in S^\#$ such that $q'x \in S$ and $qx \neq 0$. Now take any $a' \in A$, then $(a' + aq')x \in A$ and $qx \neq 0$. Thus $A < A + aQ(Q_S)$, and so $A + aQ \subset A^c = A$, hence $aQ \subset A$. Therefore A is a right ideal. To see that it is closed, assume $A < B(Q_Q)$. Then also $A < B(Q_S)$, by 1.7 (†), hence $B \subset A^c = A$, as required.

4.5. As has also been observed by Johnson, Theorem 4.3 partly generalizes a recent result of Goldie (**8**). Goldie obtained the conclusion of Theorem 4.3 (even using the classical ring of quotients) for prime rings satisfying the following ascending chain conditions as well as their symmetric duals:

(1r) Every direct sum of non-zero right ideals of S has a finite number of terms.

(2l) The ascending chain condition holds for the annihilator left ideals of S.

It is not difficult to show that the assumption of Theorem 4.3 for a prime ring S is implied by (1r) and (2l), or even by (1r) and (2r), the symmetric dual of (2l) (**15**, Propriété 12). In this connection we shall only establish one lemma.

4.6. A ring without non-zero, nilpotent ideals is called *semi-prime*. Clearly, every prime ring is semi-prime.

Lemma. *If S is any semi-prime ring satisfying* (2l), *then* $J(S_S) = 0$.

Proof. Let $\{L_i \mid i \in I\}$ be the set of all closed large right ideals of S and consider

$$J = \sum_{i \in I} L_i^l.$$

We have

$$J^{rl} = \Big(\sum_{i \in I} L_i{}^l\Big)^{rl} = \Big(\sum_{i \in F} L_i{}^l\Big)^{rl},$$

for a finite subset F of I, by (2*l*). Therefore

$$J^r = J^{rlr} = \Big(\bigcap_{i \in F} L_i{}^{lr}\Big)^{lr} = \bigcap_{i \in F} L_i{}^{lr},$$

since $A \to A^{lr}$ is also a closure operation on the lattice of right ideals of S, and the intersection of "closed" right ideals is "closed." Now a finite intersection of large right ideals is large, hence $L = J^r$ is a large right ideal. Thus $(J \cap L)^2 \subset JL = 0$. Since S is semi-prime, $J \cap L = 0$, hence $J = 0$.

Thus $L_i{}^l = 0$, for all closed, large right ideals L_i of S. This easily implies that $L'^l = 0$, for any large right ideal L' of S, as was to be shown.

5. Rational completions of bimodules.

5.1. If R and S are associative rings, a *bimodule* ${}_RM_S$ consists of a right module M_S and a left module ${}_RM$ with the same additive group such that

$$(rm)s = r(ms) \qquad (r \in R,\ m \in M,\ s \in S).$$

By a standard trick, ${}_RM_S$ may be regarded as a right module, even a unitary right module M_T. Thus let R' be anti-isomorphic with R, then we put $T = S^\# \otimes_Z R'^\#$ and write

$$m(x \otimes y') = ymx \qquad (m \in M,\ x \in S^\#,\ y \in R^\#).$$

In view of this identification it is clear that, for R- S-bimodules A, B, and C, $A < B({}_RC_S)$ must mean that ${}_RA_S$ is a submodule of ${}_RB_S$ and $\mathrm{Hom}_{R,S}(E, C) = 0$, for every submodule ${}_RE_S$ of ${}_RB_S - {}_RA_S$. We can also speak of the rational completion ${}_R\bar{M}_S$ of ${}_RM_S$, meaning that $\bar{M}_T$ is the rational completion of M_T.

5.2. Theorem. *Let ${}_RM_S$ be any bimodule, ${}_R\bar{M}_S$ its rational completion. Then the rational completions of M_S and ${}_RM$ are also bimodules ${}_R\vec{M}_S$ and ${}_R\overleftarrow{M}_S$ respectively. They are isomorphic over ${}_RM_S$ to unique submodules of ${}_R\bar{M}_S$ and will be identified with these. Their intersection ${}_R\hat{M}_S$ in ${}_R\bar{M}_S$ is the largest extension of ${}_RM_S$ satisfying $M < \hat{M}(\hat{M}_S)$ and $M < \hat{M}({}_R\hat{M})$. ${}_R\hat{M}_S$ is "b-complete" in the following sense:*

If ${}_RA_S$ and ${}_RB_S$ are any bimodules such that $A < B(\hat{M}_S)$ and $A < B({}_R\hat{M})$, then any element of $\mathrm{Hom}_{R,S}(A, \hat{M})$ can be extended to a unique element of $\mathrm{Hom}_{R,S}(B, \hat{M})$.

Proof. Every $r \in R$ determines an element of $\mathrm{Hom}_S(M, M)$, namely the map $m \to rm$, $m \in M$. Since M_S is dense in $\vec{M}_S$, this map may be extended to a unique element of $\mathrm{Hom}_S(\vec{M}, \vec{M})$, by 1.3. We may as well write this map $n \to rn$, $n \in \vec{M}$. Thus $\vec{M}$ is also an R- S-bimodule.

Now $M < \vec{M}(\vec{M}_S)$, *a fortiori* $M < \vec{M}({}_R\vec{M}_S)$. By 1.3, the injection of M into $\vec{M}$ can be extended to a unique element of $\mathrm{Hom}_{R,S}(\vec{M}, \bar{M})$, and this is easily seen to be a monomorphism.

We may identify $\vec{M}$ with its isomorphic image in $\bar{M}$. Similarly $\overleftarrow{M}$ may be regarded as a submodule of $\bar{M}$. Put $\hat{M} = \vec{M} \cap \overleftarrow{M}$. From $M < \vec{M}(\vec{M}_S)$ we immediately deduce that $M < \hat{M}(\hat{M}_S)$. By symmetry, we have also $M < \hat{M}({}_R\hat{M})$. We defer the proof that $\hat{M}$ is the largest extension of M with these two properties.

Now let $A < B(\hat{M}_S)$, $A < B({}_R\hat{M})$, and $\phi \in \mathrm{Hom}_{R,S}(A, \hat{M})$. *A fortiori*, $\phi \in \mathrm{Hom}_S(A, \hat{M})$, hence it may be extended to a unique $\vec{\phi} \in \mathrm{Hom}_S(B, \vec{M})$. Take any $r \in R$ and compare $r\vec{\phi}$ with $\vec{\phi}r \in \mathrm{Hom}_S(B, \vec{M})$. These two maps coincide on A, hence on B, since $A < B(\vec{M}_S)$. (This last statement follows from $A < B(\hat{M}_S)$ and $\hat{M} < \vec{M}(\vec{M}_S)$ by P5, where the second statement follows from $M < \vec{M}(\vec{M}_S)$ by P3.)

Thus $\vec{\phi} \in \mathrm{Hom}_{R,S}(B, \vec{M})$. In the same way, we extend ϕ to a unique $\overleftarrow{\phi} \in \mathrm{Hom}_{R,S}(B, \overleftarrow{M})$. Now both $\vec{\phi}$ and $\overleftarrow{\phi}$ may be regarded as elements of $\mathrm{Hom}_{R,S}(B, \bar{M})$. They agree on A, hence on B, since $A < B({}_R\bar{M}_S)$. (This last statement follows by P5 from $A < B({}_R\hat{M}_S)$, which is a trivial consequence of $A < B(\hat{M}_S)$, and $\hat{M} < \bar{M}({}_R\bar{M}_S)$, an immediate consequence of $M < \bar{M}({}_R\bar{M}_S)$.)

Now the image of $\vec{\phi}$ lies in $\vec{M}$, the image of $\overleftarrow{\phi}$ in $\overleftarrow{M}$, hence their common image lies in $\vec{M} \cap \overleftarrow{M} = \hat{M}$, and so we obtain an element of $\mathrm{Hom}_{R,S}(B, \hat{M})$.

Finally, assume that $M < N(N_S)$ and $M < N({}_RN)$. It follows by a standard argument that N may be regarded as a unique submodule of $\hat{M}$. (Indeed, since $N < \hat{M}(\hat{M}_S)$, P5 yields $M < N(\hat{M}_S)$, and similarly $M < N({}_R\hat{M})$. In view of the completeness property just proved, the injection of M into $\hat{M}$ may be extended to a unique element of $\mathrm{Hom}_{R,S}(N, \hat{M})$, and this is easily seen to be a monomorphism.) Thus $\hat{M}$ is, up to isomorphism, the largest bimodule N with the prescribed properties.

5.3. THEOREM. *Let S be a ring, ${}_S\bar{S}_S$ its rational completion as a bimodule. The maximal rings $\vec{S}$ and $\overleftarrow{S}$ of right and left quotients of S, regarded as S-S-bimodules, are isomorphic to unique submodules of ${}_S\bar{S}_S$, and will be identified with these. Their intersection $\hat{S}$ is a subring of both. It is the largest ring extension of S which is both a right and a left ring of quotients of S.*

$\hat{S}$ is the maximal ring of right and left quotients of S of Johnson and Wong

(**19**, 8). A special case had previously been studied by Utumi (**17**, 5.3). The present construction is more symmetrical than these earlier ones.

Proof. All of this follows immediately from 5.2, with the exception of the fact that the operations of multiplication in the rings $\overrightarrow{S}$ and $\overleftarrow{S}$ coincide on their intersection $\hat{S}$.

As was shown in (**5**), $\overrightarrow{S}$ is a ring with multiplication $*$ (say) such that $q * s = qs$, for all $q \in \overrightarrow{S}$ and $s \in S$. By symmetry, $\overleftarrow{S}$ is a ring with multiplication $\circ$ (say), such that $s \circ p = sp$, for all $s \in S$ and $p \in \overleftarrow{S}$. We wish to show that $p \circ q = p * q \in \hat{S}$, for all p and q in $\hat{S}$.

Let us write

$$Y = \{y \in S^\#|\ qy \in S\}, \qquad X = \{x \in S^\#|\ xp \in S\}.$$

A simple calculation show sthat

$$x(p \circ q)y = (xp)(qy) = x(p * q)y \qquad (x \in X,\ y \in Y),$$

and so $X(p \circ q - p * q)Y = 0$. Since

$$S^\# - Y \cong (qS^\# + S) - S,$$

we deduce from P3 that $Y < S^\#(S_S)$, and similarly that $X < S^\#({}_SS)$. The result now follows if, for any $m \in \bar{S}$, $XmY = 0$ implies $m = 0$. In view of the representation of bimodules as unitary right modules (see 5.1), this may be inferred from the following lemma.

5.4. Lemma. *If M_R, N_S, A_R, B_S and $C_{R\otimes S}$ are right modules such that $A < M(C_R)$ and $B < N(C_S)$ then $[A \otimes B] < M \otimes N(C_{R\otimes S})$.*

Here $[A \otimes B]$ is the set of all

$$\sum_{i=1}^{k} a_i \otimes b_i \in M \otimes N$$

with $a_i \in A$ and $b_i \in B$.

Proof. Let D be any $R \otimes S$-submodule of $M \otimes N$ and consider $\phi \in \mathrm{Hom}_{R\otimes S}(D, C)$ such that $[A \otimes B] \subset \ker \phi$. We wish to show that $\mathrm{im}\ \phi = 0$.

Take a k-tuple $(a_1, \ldots, a_k)$ of elements of A. Let D' be the set of all k-tuples $(n_1, \ldots, n_k)$ of elements of N such that

$$\sum_{i=1}^{k} a_i \otimes n_i \in D.$$

Clearly, D' is an S-submodule of $kN = N \oplus \ldots \oplus N$. Let

$$\phi'(n_1, \ldots, n_k) = \phi\left(\sum_{i=1}^{k} a_i \otimes n_i\right),$$

then $\phi' \in \mathrm{Hom}_S(D', C)$ and $kB \subset \ker \phi'$. Now, $B < N(C_S)$, hence $0 < N - B(C_S)$, and therefore $0 < k(N - B)(C_S)$. But $k(N - B) \cong kN - kB$, hence $kB < kN(C_S)$. Therefore $\mathrm{im}\, \phi' = 0$, and so $[A \otimes N] \subset \ker \phi$. Repeating the whole argument on the other side, we finally obtain $M \otimes N \subset \ker \phi$, as required.

5.5. Let D be a skew-field, ${}_DV$ and V'_D left and right D-modules (vector spaces) respectively. Put

$$B = \mathrm{Hom}_{D,D}(V \otimes_Z V', D);$$

this is clearly an additive group. We may regard B as the module of *bilinear forms* from $V \times V'$ into D. There is a canonical isomorphism

$$B \cong \mathrm{Hom}_D(V, \mathrm{Hom}_D(V', D)).$$

Thus, for any $b \in B$ and $v \in V$, we may regard vb as an element of $\mathrm{Hom}_D(V', D)$ such that

$$(vb)v' = vbv' \qquad (v' \in V').$$

(We write vbv' in place of $b(v \otimes v')$.)

An element b_0 of B is called *non-degenerate* if

$$vb_0 = 0 \Rightarrow v = 0 \quad \text{and} \quad b_0v' = 0 \Rightarrow v' = 0 \quad (v \in V, v' \in V').$$

If B contains a non-degenerate element b_0, V and V' are called *dual vector spaces* (**9**, p. 69).

Put $S = V' \otimes_D V$. This is turned into a ring with an obvious multiplication, as illustrated by

$$(v_1' \otimes v_1)(v_2' \otimes v_2) = v_1' \otimes (v_1 b_0 v_2')v_2.$$

Moreover, one obtains in a natural way the bimodules ${}_DV_S$, ${}_SV'_D$, and ${}_SB_S$.

If V'_D and ${}_DV$ are dual vector spaces, the mapping $v \to vb_0$ is a monomorphism of ${}_DV$ into $\mathrm{Hom}_D(V', D)$, and this induces a monomorphism of $\mathrm{Hom}_D(V, V)$ into B, its image being $\{b \in B \mid Vb \subset Vb_0\}$. We also have an isomorphic embedding of ${}_SS_S$ into ${}_SB_S$. Indeed, the element

$$s = \sum_{i=1}^{n} v_i' \otimes v_i$$

of S gives rise to the bilinear form (s) where

$$v(s)v' = \sum_{i=1}^{n} (vb_0v_i')(v_ib_0v').$$

THEOREM. *Let V'_D and ${}_DV$ be dual vector spaces with a non-degenerate bilinear form b_0, and let S be the ring $V' \otimes_D V$. Then the bimodule ${}_SB_S$ of bilinear forms from $V \times V'$ into D is a rational extension of ${}_SS_S$, and the maximal rings of right quotients, left quotients, and right and left quotients of S may be realized thus:*

$$\overrightarrow{S} = \{b \in B \mid bV' \subset b_0 V'\},$$
$$\overleftarrow{S} = \{b \in B \mid Vb \subset Vb_0\},$$
$$\hat{S} = \{b \in B \mid bV' \subset b_0 V' \text{ and } Vb \subset Vb_0\}.$$

We omit the proof of this theorem, which is another formulation of Utumi's results (**17**, 5.1, 5.3), in the hope of improving it at a later time in two directions: to identify the rational completion of ${}_S S_S$ and to extend the result to projective modules over prime rings.

5.6. The preceding theorem may be applied to obtain $\hat{S}$ for any prime ring S with non-zero socle. As is well known (**16**), such a ring is a primitive ring, hence we may apply Utumi's result (see 1.5). Thus $\overleftarrow{S} \cong \mathrm{Hom}_D(V, V)$, where $D = eSe$ is a skew-field and $V = eS$ is a left D-module. Dually, also $\overrightarrow{S} \cong \mathrm{Hom}_D(V', V')$, where $V' = Se$ is a right D-module. Utumi also computed $\hat{S}$ (**17**, 5.3), but a more symmetric form of $\hat{S}$ may be obtained by 5.5.

Indeed, it is well known (**9**, page 77) that ${}_D V$ and V'_D are dual vector spaces. One easily verifies that S is a ring of right and left quotients of SeS, the latter being isomorphic to $V' \otimes_D V = S_0$, say. Thus $\hat{S} = \hat{S}_0$, and this is determined by 5.5.

6. On semi-prime rings.

6.1. With any bimodule ${}_R M_S$ we may associate the lattices $L({}_R M_S)$, $L({}_R M)$, and $L(M_S)$. In addition, we shall be interested in the lattice $L^b({}_R M_S)$, which consists of all *b-closed* submodules of ${}_R M_S$, where b is a closure operation defined on the lattice of all submodules of ${}_R M_S$ as follows:

Let ${}_R A_S$ be any submodule of ${}_R M_S$ then ${}_R A^b{}_S$ is the largest submodule ${}_R B_S$ of ${}_R M_S$ such that

$$(\ddagger) \qquad A < B(M_S) \quad \text{and} \quad A < B({}_R M).$$

We will show that ${}_R A^b{}_S$ is in fact the intersection of the closure of A_S in M_S with the closure of ${}_R A$ in ${}_R M$.

Indeed, let the closure operation for submodules of M_S be denoted by c. Take any element r of R, then

$$r(A^c) \subset (rA)^c \subset A^c,$$

by C3 and the fact that A is a left R-module. Thus we have a bimodule ${}_R A^c{}_S$. In the same way, if the closure operation for submodules of ${}_R M$ is denoted by d, we obtain a bimodule ${}_R A^d{}_S$. Put $B = A^c \cap A^d$, then B is a bimodule and ($\ddagger$) holds.

On the other hand, assume that ${}_R B_S$ is any submodule of ${}_R M_S$ satisfying ($\ddagger$). Then $B \subset A^c$ and $B \subset A^d$, hence $B \subset A^c \cap A^d$, as was to be shown.

Henceforth we write $A^b = A^c \cap A^d$.

6.2. We now make the blanket assertion:

All results obtained for $L(M_S)$ in § 2 and § 3 remain valid for $L^b({}_RM_S)$, *mutatis mutandis.*

Indeed, the results in § 2 and § 3 were based only on these facts: the existence of a closure operation, the existence of a rational completion, and properties P0 to P5 for the ternary relation $A < B(C_S)$ among right modules.

Since we have already established the existence of a b-closure and a b-completion, it remains to verify properties P0 to P5 for the ternary relation

$$A < B(C_S) \quad \text{and} \quad A < B({}_RC)$$

among bimodules. This is a routine verification. For example, P5 asserts for bimodules that

$$[A < B(C_S) \text{ and } A < B({}_RC) \text{ and } C < D(D_S) \text{ and } C < D({}_RD)]$$
$$\Rightarrow [A < B(D_S) \text{ and } A < B({}_RD)].$$

This implication clearly follows from the separate implications for left modules and right modules.

From the above blanket assertion we must except the special construction in 2.1.

In translating results from one situation to the other, we must make the following replacements:

Replace	c	by	b,
	$L(M)$	by	$L^b(M)$,
	closed	by	*b-closed,*
	dense	by	*b-dense,*
	rationally complete	by	*b-complete,*
	rational extension	by	*right and left rational extension.*

Here a submodule ${}_RA_S$ of ${}_RM_S$ is called *b-dense* if $A^b = M$.

In future, the analogue of (let us say) 2.5 for b-closure will be denoted by 2.5^b.

6.3. If S is a ring, we are particularly interested in $L^b({}_SS_S)$, which we shall denote more briefly by $L^b(S)$.

PROPOSITION. *For an associative ring S, $L^b(S)$ has at most two elements if and only if either S is a prime ring or $S^2 = 0$ and the additive group of S is cyclic of prime order or a subgroup of the additive group of rationals.*

Proof. We proceed in three steps.

(1) If S is a non-zero prime ring, then $L^b(S)$ has exactly two elements.

Indeed, let A be any non-zero ideal, then $A^l = 0$. From this one easily deduces that A_S is dense in S_S. See, for example (**5**, 6.4). Similarly ${}_SA$ is dense in ${}_SS$, hence A is b-dense in S.

(2) If $L^b(S)$ has at most two elements and $S^2 \neq 0$, then S is a prime ring.

Indeed, assume that every non-zero ideal is b-dense and $S^2 \neq 0$. Suppose $S^l \neq 0$, then $sS = 0$, for some $s \neq 0$. Then $S^\# s$ is a non-zero ideal, hence it is b-dense in S. But $S^\# sS = 0$, hence $SS = 0$, contrary to assumption. Thus $S^l = 0$, and similarly $S^r = 0$.

Now suppose $sSs' = 0$, $s' \neq 0$. Since $S^r = 0$, $Ss' \neq 0$. Since, $S^l = 0$, $Ss'S \neq 0$. Thus $Ss'S$ is b-dense in S. But $sSs'S = 0$, hence $sS = 0$. Since $S^l = 0$, we have $s = 0$. Therefore S is a prime ring.

(3) If $S^2 = 0$, then $L^b(S) = L^b({}_Z S_Z) = L(S)$, where Z is the ring of integers. The result now follows from 3.6.

6.4. We may also ask when $L^b(S)$ is a complemented lattice. Essentially, this implies that $\hat{S}$ is a semi-prime ring and that $L^b(S)$ is a Boolean algebra, as we shall see.

Porposition. *If S is a ring for which $L^b(S)$ is complemented, then $L^b(S) \cong L^b(\hat{S})$, where $\hat{S}$ is the maximal ring of right and left quotients of S. If furthermore $S^l = 0$, then $\hat{S}$ is semi-prime.*

Proof. Clearly, ${}_S S_S$ is b-dense in ${}_S \hat{S}_S$. Hence, by 2.5^b, $L^b(S) \cong L^b({}_S \hat{S}_S)$. We claim that the latter is actually $L^b({}_{\hat{S}} \hat{S}_{\hat{S}}) = L^b(\hat{S})$. Indeed, this will follow from the lemma below, which asserts that all b-closed submodules of ${}_S \hat{S}_S$ are b-closed ideals in $\hat{S}$.

If $S^l = 0$ then $\hat{S}^l \cap S = 0$, hence, also $\hat{S}^l = 0$, since $S < \hat{S}(\hat{S}_S)$. For the remainder of the proof we may as well assume that $\hat{S} = S$ and $S^l = 0$. Suppose that A is a non-zero, nilpotent ideal of S, say $A^k = 0$ and $A^{k-1} \neq 0$, for $k > 2$. Let $B = A^{k-1}$ and consider its b-closure B^b. Now $B < B^b(S_S)$, $BB^b \subset S$, and $B^2 = 0$, hence $BB^b = 0$. Applying 3.3^b, we obtain $S = B^b \oplus C$, where C is another ideal of S. Therefore $BC \subset B^b C = 0$, hence $BS = BB^b + BC = 0$. Since $S^l = 0$, we deduce $B = 0$, a contradiction. Thus S contains no non-zero, nilpotent ideal, and so is semi-prime.

6.5. Lemma. *If S is a ring such that $L^b(S)$ is complemented, $\hat{S}$ its maximal ring of right and left quotients, then any b-closed submodule of ${}_S \hat{S}_S$ is a b-closed ideal of $\hat{S}$.*

Proof. Let $A \in L^b({}_S \hat{S}_S)$. By 3.3^b, $\hat{S} = A + B$, $A \cap B = 0$. Now $A \cap S$ and $B \cap S$ are ideals of S, and

$$(A \cap S)(B \cap S) \subset A \cap B = 0.$$

By 2.5^b, the b-closure of $A \cap S$ in ${}_S \hat{S}_S$ is A and that of $B \cap S$ is B.

Take any element a of $A \cap S$, then $aB \subset S$, $a(B \cap S) = 0$ and $B \cap S < B(\hat{S}_S)$, hence $aB = 0$. Thus $(A \cap S)B = 0$. Arguing similarly on the other side, we obtain $AB = 0$. By symmetry also $BA = 0$, and so A and B are ideals of $\hat{S}$.

Let A' be the b-closure of A in ${}_{\hat{S}} \hat{S}_{\hat{S}}$, then $A < A'(\hat{S}_{\hat{S}})$. But $S < \hat{S}(\hat{S}_S)$, hence $A < A'(\hat{S}_S)$, by 1.7(†). Similarly $A < A'({}_S \hat{S})$, and so A' is contained in the b-closure of A in ${}_S \hat{S}_S$, which is just A. Thus $A' = A$, as required.

6.6. Johnson has shown in **(12, II)**—among many other interesting results—that the annihilator ideals of a semi-prime ring form a complete Boolean algebra. This is also contained in the following:

THEOREM. *If S is a semi-prime ring, then $L^b(S)$ is a complete Boolean algebra, whose elements are the annihilator ideals of S. If $\hat{S}$ is the maximal ring of right and left quotients of S, then $\hat{S}$ is also semi-prime, $L^b(\hat{S}) \cong L^b(S)$, and the elements of $L^b(\hat{S})$ are the direct summands of $\hat{S}$.*

Proof. Let S be semi-prime. We first verify the following condition:

(*) For each ideal A of S there exists an ideal A^* such that, for any ideal B of M, $A \cap B = 0$ if and only if $B \subset A^*$.

Indeed, let A^r be the right annihilator of A in S, then $(A \cap A^r)^2 \subset AA^r = 0$, and so $A \cap A^r = 0$, since S is semi-prime. If B is any ideal such that $A \cap B = 0$, then $AB = 0$, hence $B \subset A^r$. Thus the condition holds with $A^* = A^r$.

The following consequences of (*) are immediate:

(1) A^* is uniquely determined.
(2) $A \subset A^{**}$, $A^{***} \subset A^*$, $A \subset B \Rightarrow B^* \subset A^*$.
(3) $A \to A^{**}$ is a closure operation.
(3) A^{**} is the largest ideal of S in which A is a large S-S-submodule.
(5) For any collection $\{A_i\}_{i \in I}$ of ideals of S,

$$\Big(\sum_{i \in I} A_i\Big)^* = \bigcap_{i \in I} A_i^*.$$

(6) The ideals A of S such that $A^{**} = A$ form a complete Boolean algebra with set intersection as meet and * as complementation. The join of a family $\{A_i\}_{i \in I}$ of elements of this Boolean algebra is given by

$$\bigvee_{i \in I} A_i = \Big(\sum_{i \in I} A_i\Big)^{**} = \Big(\bigcap_{i \in I} A_i^*\Big)^*.$$

We omit the straightforward derivations of (1) to (6) from (*). Since we could also have taken $A^* = A^l$, the left annihilator of A in S, it follows from (1) that $A^* = A^r = A^l$. Thus the right annihilator ideals of S are the same as the left annihilator ideals.

Next, we shall show that

$$(**) \qquad A < A^{**}(S_S).$$

Take $x \in A^{**}$, $0 \neq s \in S$, we seek $y \in S^\#$ such that $sy \neq 0$ and $xy \in A$. In fact, we shall find y in S. We have apparently three cases:

Case 1. $sA \neq 0$. Take $y \in A$ such that $sy \neq 0$. Then $xy \in SA \subset A$.

Case 2. $sA^* \neq 0$. Take $y \in A^*$ such that $sy \neq 0$. Then $xy \in A^{**}A^* = 0 \subset A$.

Case 3. $sA = 0$ and $sA^* = 0$. Then $s(A + A^*) = 0$, hence $s \in (A + A^*)^* = A^* \cap A^{**} = 0$. Since $s \neq 0$, this case does not really arise.

Now A^{**} is a closed submodule of S_S, by 2.7 (1). Hence, by (**), it is the closure of A in S_S. By symmetry, it is also the closure of A in ${}_SS$, hence it is the b-closure of A in ${}_SS_S$. Thus the annihilator ideals coincide with the b-closed ideals, and so $L^b(S)$ is a complete Boolean algebra, by (6) above.

It follows from 6.4 that $L^b(S) \cong L^b(\hat{S})$ and that $\hat{S}$ is semi-prime. Hence the annihilator ideals of S are also the b-closed ideals of S, and these are the direct summands of $\hat{S}$, in view of 3.3^b. The proof is now complete.

6.7. The *Dedekind–MacNeille* completion of a partially ordered set S is a complete lattice, whose elements are the subsets of S of the form (see (**1**, p. 58, Theorem 12)): the set of all lower bounds of the set of all upper bounds of a non-empty subset K of S. The following corollary to Theorem 6.6 contains a new proof of the main result of (**2**) for Boolean rings with 1.

COROLLARY. *The Dedekind–MacNeille completion of a Boolean ring with* 1 *is given by*

$$L^b(S) \cong L^b(\hat{S}) \cong \hat{S} = \vec{S}.$$

Proof. Let K be any non-empty subset of S, then the set of its upper bounds is

$$K' = \{s \in S \mid \forall_{k \in K}\, sk = k\} = \{s \in S \mid 1 - s \in K^*\},$$

and the set of all lower bounds of K' is

$$\{t \in S \mid \forall_{s \in K'}\, st = t\} = \{t \in S \mid K^*t = 0\} = K^{**}.$$

Thus the Dedekind–MacNeille completion of S consists precisely of the annihilator ideals of S, hence coincides with $L^b(S) \cong L^b(\hat{S})$, by Theorem 6.6. Now it is easily verified that $\hat{S} = \vec{S}$ is a Boolean ring (see (**2**, Corollary 2)). Therefore $L^b(\hat{S}) \cong \hat{S}$, by the last part of Theorem 6.6.

6.8. We have called a ring *semi-prime* if it has no non-zero, nilpotent ideals. It is known (see, for example (**9**, p. 196)) that a semi-prime ring may also be characterized as a ring in which the intersection of all prime ideals is 0. A *prime ideal* of S is any ideal P such that $S - P$ is a prime ring. The following two assertions are equivalent characterizations of prime ideals:

(a) For all ideals A and B, if $AB \subset P$ then $A \subset P$ or $B \subset P$.

(b) For all elements s and s', if $sSs' \subset P$, then $s \in P$ or $s' \in P$.

In what follows, $\mathscr{P}(S)$ will denote the set of proper prime ideals of S.

It is easily verified that, for any ideal A of a semi-prime ring S,

$$A^* = \bigcap\{P \in \mathscr{P}(S) \mid A \not\subset P\}.$$

This was used in a different approach to Theorem 6.6 by the author in (Amer. Math. Soc. Notices, 7 (1960), p. 92). It turns out that every proper prime ideal P is either b-closed ($P^{**} = P$) or b-dense ($P^{**} = S$) in S. The former are also the maximal proper b-closed ideals of S.

Condition (*), which is responsible for associating a Boolean algebra with the ring S, might also hold for rings which are not semi-prime. It can be shown that (*) is in fact equivalent to the vanishing of the intersection of all ideals P' of S such that

$$A \cap B = 0 \Rightarrow \text{ either } A \subset P' \text{ or } B \subset P',$$

for any ideals A and B. A similar result holds for modules, but it would take us too far afield to go into further details here.

6.9. As was pointed out by McCoy **(16)**, the set $\mathscr{P}(S)$ of proper prime ideals of a semi-prime ring becomes a topological space under the usual Stone topology, the open sets being precisely the sets

$$\Gamma A = \{P \in \mathscr{P}(S) | \, A \not\subset P\},$$

where A is any ideal of S.

If V is any open subset of $\mathscr{P}(S)$, we introduce the ideal

$$\Delta V = \cap_{P \notin V} P.$$

Then $\Delta\Gamma A = A^*$, the annihilator of A. On the other hand, $\Gamma\Delta V = V^\perp$ is easily seen to be the interior of the complement of V, also called the *exterior* of V. A set of the form $V^\perp$ is called a *regular open* set. The open set U is regular open if and only if $(U^\perp)^\perp = U$.

THEOREM. *If S is a semi-prime ring, the mapping $A \to \Gamma A$ is an isomorphism of the complete Boolean algebra of annihilator ideals of S onto the algebra of regular open sets in the prime ideal space $\mathscr{P}(S)$.*

Proof. One easily verifies that $\Gamma(A^*) = (\Gamma A)^\perp$ and $\Gamma(A \cap B) = \Gamma(A) \cap \Gamma(B)$, for annihilator ideals A and B. Thus Γ is a lattice homomorphism. Now $\Delta\Gamma\Delta$ is the inverse mapping of Γ; for let A be any annihilator ideal, V any regular open set, then

$$(\Delta\Gamma\Delta)\Gamma A = (A^*)^* = A, \qquad \Gamma(\Delta\Gamma\Delta)\, V = (V^\perp)^\perp = V.$$

The analogous result for maximal ideal spaces of commutative semi-simple rings with 1 was recently obtained by Fine, Gillman, and the present author. The proofs of these two results are practically identical.

7. On the structure of semi-prime rings. We wish to present some results on the structure of semi-prime rings, which resemble those of Dieudonné **(3)**. We require three lemmas which we have stated together for convenience.

7.1. LEMMA. *Let $S = C \oplus D$ as a direct sum of rings.*

(1) *If S is a b-complete ring, then so is C.*

(2) $L^b({}_S C_S) = L^b(C)$.

(3) *${}_S C_S$ is a b-complete bimodule, if C is a b-complete ring, and $S^l = 0 = S^r$.*

Proof.

(1) Let $A < B(C_C)$, $A < B({}_CC)$, and $\phi \in \mathrm{Hom}_{C,C}(A, C)$. We may turn A and B into S-S-bimodules by demanding that $DB = 0$ and $BD = 0$. Then also $A < B(C_S)$ and $A < B({}_SC)$. Now ϕ may be regarded as an element of $\mathrm{Hom}_{S,S}(A, S)$, hence it may be extended to an element ψ of $\mathrm{Hom}_{S,S}(B, C)$. Then $\pi\psi \in \mathrm{Hom}_{C,C}(B, S)$ extends ϕ, where π is the projection $C \oplus D \to C$.

(2) If $A \in L^b({}_SC_S)$, a striaghtforward argument shows that $A \in L^b(C)$. The converse is a bit more difficult: Let A be a b-closed ideal in C, B its b-closure in ${}_SC_S$. Then $A < B(C_S)$, and so, for any $b \in B$ and $0 \neq c \in C$, we can find $x \in S^\#$ such that $bx \in A$ and $cx \neq 0$. Now $x = c + d + z$, where $c \in C$, $d \in D$, and z is an integer. Since $bd = 0$ and $cd = 0$, we may as well take $d = 0$, so that $x \in C^\#$. Thus $A < B(C_C)$ and, by symmetry, $A < B({}_CC)$. Since A was a b-closed ideal of C, we have $A = B$, and so A is also b-closed in ${}_SC_S$, as required.

(3) Let $A < B(C_S)$, $A < B({}_SC)$, and $\phi \in \mathrm{Hom}_{S,S}(A, C)$. Let $0 \neq c \in C$ and $b \in B$, we can find $x \in S^\#$ such that $cx \neq 0$ and $bx \in A$. Now $S^l = 0$, hence $cxC = cxS \neq 0$, and so there exists $c' \in C$ such that $cxc' \neq 0$. But $bxc' \in A$ and $xc' \in C$, hence $A < B(C_C)$. Similarly $A < B({}_CC)$. Since C is b-complete, ϕ can be extended to $\psi \in \mathrm{Hom}_{C,C}(B, C)$. We will show that $\psi \in \mathrm{Hom}_{S,S}(B, C)$.

Indeed, it suffices to show that $\psi(db) = d(\psi b)$, for any $b \in B$ and $d \in D$. Given d, the mapping $b \to \psi(db)$ belongs to $\mathrm{Hom}_C(B, C)$, and $\psi(dA) = \phi(dA) = d(\phi A) = 0$. Now we recall that $A < B(C_C)$, hence $\psi(dB) = 0$. Since also $d(\psi B) \subset dC = 0$, the result follows.

7.2. Proposition. *If S is the weak direct sum of the set of rings $\{S_i\}_{i\in I}$ and $S^l = 0 = S^r$, then its maximal ring of right and left quotients $\hat{S}$ is the complete direct sum of the $\hat{S}_i$.*

This is the two-sided analogue of **(17, 2.1)**.

Proof. We shall prove this in three steps.

(1) A complete direct sum of b-complete bimodules is b-complete.

Indeed, let

$$M = \prod_{i\in I} M_i$$

where the M_i are b-complete R-S-modules. Suppose $A < B(M_S)$ and $A < B({}_RM)$. Let $\phi \in \mathrm{Hom}_{R,S}(A, M)$. Since there is a well-known monomorphism of M_i into M, we have $A < B(M_{iS})$ and $A < B({}_RM_i)$. Now let π_i be the canonical epimorphism of M onto M_i, then $\pi_i\phi \in \mathrm{Hom}_{R,S}(A, M_i)$. By b-completeness of M_i, this may be extended to $\psi_i \in \mathrm{Hom}_{R,S}(B, M_i)$. By definition of direct products, there exists a unique $\psi \in \mathrm{Hom}_{R,S}(B, M)$ such that $\pi_i\psi = \psi_i$, for all $i \in I$. Now, for any $a \in A$, $\pi_i(\psi a) = \psi_i a = \pi_i(\phi a)$, hence $\psi a = \phi a$, and so ψ extends ϕ.

(2) If

$$S = \dot{\sum_{i\in I}} S_i$$

is the weak direct sum of rings S_i, and T_i is a ring of right quotients of S_i, then

$$T = \prod_{i \in I} T_i$$

is a ring of right quotients of S.

Actually we only require the known case $S^l = 0$. In the general case one might proceed thus:

Let $t \in T$, $s \in S$. Denoting by t_i the ith component of t, then $(ts)_i = t_i s_i \in T_i S_i \subset T_i$, hence T is a right S-module. We claim that $S < T(T_S)$. Indeed, let $t' \neq 0$ and $t \in T$, we seek $x \in S$ such that $t'x \neq 0$ and $tx \in S$.

Since $t' \neq 0$, there exists $k \in I$ such that $t_k' \neq 0$. Now T_k is a ring of right quotients of S_k, hence we can find $x_k \in S_k^\#$ such that $t_k' x_k \neq 0$ and $t_k x_k \in S_k$. Putting $x_i = 0$ for $i \neq k$, we obtain an element x of $S^\#$, for which it is easily verified that $t'x \neq 0$ and $tx \in S$.

(3) We now prove the proposition. By (2) and symmetry,

$$T = \prod_{i \in I} \hat{S}_i$$

is a ring of right and left quotients of

$$S = \dot{\sum_{i \in I}} S_i.$$

Now each of the $\hat{S}_i$ is b-complete as a ring, hence also as a T-T-bimodule, by Lemma 7.1. Therefore, by (1), T is also b-complete. Now $T < \hat{S}(T_S)$ and $S < T(T_S)$, hence $T < \hat{S}(T_T)$, by 1.7 (†). By symmetry also $T < \hat{S}({}_T T)$, hence $T = \hat{S}$.

7.3. We recall that a ring is called *b-complete* if it coincides with its maximal ring of right and left quotients.

THEOREM. *If S is a b-complete semi-prime ring, then $S = C \oplus C^*$, where C is the socle of $L^b(S)$. Let $\{A_i\}_{i \in I}$ be the set of all atoms of $L^b(S)$, then*

$$C \cong \prod_{i \in I} A_i, \qquad C^* = \bigcap_{i \in I} A_i^*.$$

The A_i are b-complete prime rings and C^* is a b-complete semi-prime ring such that $L^b(C^*)$ has no atoms.

Proof. Since $L^b(S)$ is a Boolean algebra, it is a complemented distributive lattice. By 2.9^b, the socle C of $L^b(S)$ is the b-closure of the weak direct sum of the A_i, and by 6.6, $S = C \oplus C^*$. By Lemma 7.1, C and C^* are also b-complete rings. The A_i are prime rings by 2.6 and 6.3, they are b-complete by 7.1.

Let B be the sum of the atoms of $L^b(C)$, then C is the b-closure of B in S, hence $B < C(S_S)$, and so $B < C(C_S)$. We claim that $B < C(C_B)$.

Indeed, let $c' \neq 0$ and $c, c' \in C$. Since $S^l = 0$, we have $c'S \neq 0$. But $c'C^* = 0$, hence $c'C \neq 0$. Now $B < C(C_S)$, hence $c'B \neq 0$. Thus we can pick

$b \in B$ such that $c'b \neq 0$. Since B is an ideal of C, we also have $cb \in B$, and therefore $B < C(C_B)$.

By symmetry also $B < C({}_BC)$, and so C is a ring of right and left quotients of B. Since C is b-complete, $C = \hat{B}$, the maximal ring of right and left quotients of B. By 7.2, we have

$$C \cong \prod_{i \in I} A_i.$$

We now turn our attention to C^*. We have

$$C^* = B^{***} = B^* = \Big(\sum_{i \in I} A_i\Big)^* = \bigcap_{i \in I} A_i^*.$$

As pointed out before, C^* is a b-complete ring. Suppose there is an atom A of $L^b(C^*)$. By Lemma 7.1, this lattice is the same as $L^b({}_SC^*{}_S)$. Now all b-closed submodules of ${}_SC^*{}_S$ are b-closed ideals of S, by 2.6^b. Thus A is a b-closed ideal of S.

Suppose A contains the non-zero ideal J of S, then $J < A(C^*{}_S)$, and so J is a large submodule of S, hence $A \subset J^{**}$, by (4) in the proof of 6.6. In view of 6.6, A is contained in the b-closure of J. Thus A is an atom of $L^b(S)$.

Now all atoms of $L^b(S)$ are contained in the socle C, hence $A \subset C \cap C^* = 0$, a contradiction. Therefore $L^b(C^*)$ has no atoms. To see that C^* is semi-prime, assume that N is a nilpotent ideal of C^*. Then N is also a nilpotent ideal of $S = C \oplus C^*$, and so $N = 0$, as required.

7.4. An immediate consequence of the preceding theorem is the following.

COROLLARY. *If S is a semi-prime b-complete ring whose Boolean algebra of annihilator ideals is atomic, then S is a complete direct product of b-complete prime rings.*

Theorem 7.3 reduces the study of all b-complete semi-prime rings S to three special cases.

Case 1. S is a b-complete prime ring with non-zero socle. This case is completely described in terms of dual vector spaces by 5.5 and 5.6.

Case 2. S is a b-complete prime ring with zero socle. Section 4 throws some light on prime rings with zero socle, but it is not clear whether this is helpful here.

Case 3. S is a b-complete semi-prime ring for which $L^b(S)$ has no atoms. Such rings are very common, as the following example shows.

7.5. *Example.* Let X be a compact Hausdorff space without isolated points. Following **(6)**, we consider the ring $C(X)$ of all continuous functions from X to the real line, under point-wise addition and multiplication. With every point $x \in X$ there is associated a maximal ideal $M_x = \{f \in C(X) \mid f(x) = 0\}$, and every maximal ideal has this form. Clearly $\bigcap_{x \in X} M_x = 0$, hence $C(X)$ is semi-prime (even semi-simple). Since X has no isolated points, also

$\bigcap_{x \in X - \{x_0\}} M_x = 0$, for any point $x_0 \in X$. It is known (6, 2.11) that every prime ideal P is contained in a unique maximal ideal M_P. Let $\mathscr{P}$ be the set of all prime ideals, then

$$P^* = \bigcap \{P' \in \mathscr{P} \mid P \not\subset P'\} \subset \bigcap \{M_x \neq M_P \mid x \in X\} = 0.$$

Now it is not difficult to show, for any semi-prime ring, that every maximal proper annihilator ideal has the form $P = P^{**}$, where P is a prime ideal. Here $P^{**} = 0^* = C(X)$, hence there are no maximal proper annihilator ideals. Therefore the Boolean algebra of annihilator ideals has no atoms.

7.6. One can also obtain a kind of converse to Theorem 7.3. We shall here be content to remark one (probably well-known) fact.

LEMMA. *A complete direct product of semi-prime rings is semi-prime.*

Proof. First we observe that, if $S = C \oplus D$ as a direct sum of rings, then any prime ideal P of C gives rise to a prime ideal $P + D$ of S.

Now let $\{S_i\}_{i \in I}$ be a set of semi-prime rings, S their complete direct product. Let $s \in S$ and suppose that s lies in every prime ideal of S. In view of the above observation, the component s_i of s in S_i lies in every prime ideal of S_i. Since all S_i are semi-prime, $s_i = 0$, for all $i \in I$, hence $s = 0$.

COROLLARY. *A complete direct product of b-complete semi-prime rings is a b-complete semi-prime ring.*

REFERENCES

1. G. Birkhoff, *Lattice theory* (New York, 1948).
2. B. Brainerd and J. Lambek, *On the ring of quotients of a Boolean ring*, Can. Math. Bull., *2* (1959), 25–29.
3. J. Dieudonné, *Les idéaux minimaux dans les anneaux associatifs*, Proc. Inter. Congr. Math., vol. II (1950), 44–48.
4. B. Eckmann and A. Schopf, *Über injektive Moduln*, Arch. der Math., *4* (1956), 75–78.
5. G. D. Findlay and J. Lambek, *A generalized ring of quotients I, II*, Can. Math. Bull., *1* (1958), 77–85, 155–167.
6. L. Gillman and M. Jerison, *Rings of continuous functions* (New York, 1960).
7. A. W. Goldie, *Decompositions of semi-simple rings*, J. London Math. Soc., *31* (1956), 40–48.
8. ——— *The structure of prime rings under ascending chain conditions*, Proc. London Math. Soc. (3), *8* (1958), 589–608.
9. N. Jacobson, *Structure of rings* (Providence, 1956).
10. R. E. Johnson, *The extended centralizer of a ring over a module*, Proc. Amer. Math. Soc., *2* (1951), 891–895.
11. ——— *Semi-prime rings*, Trans. Amer. Math. Soc., *76* (1954), 375–388.
12. ——— *Structure theory of faithful rings I, II*, Trans. Amer. Math. Soc., *84* (1957), 508–522, 523–544.
13. P. R. Halmos, *Boolean algebra* (mimeographed, Chicago, 1959).

14. I. Kaplanski, *Infinite abelian groups* (Ann Arbor, 1954).
15. L. Lesieur and R. Croisot, *Anneaux premiers Noethériens à gauche*, Ann. Sci. Ec. Norm. Sup. (3), *76* (1959), 161–183.
16. N. H. McCoy, *Prime ideals in general rings*, Amer. J. Math., *71* (1949), 823–833.
17. Y. Utumi, *On quotient rings*, Osaka Math. J., *8* (1956), 1–18.
18. ——— *On a theorem on modular lattices*, Proc. Japan Acad., *25* (1959), 16–21.
19. E. T. Wong and R. E. Johnson, *Self-injective rings*, Can. Math. Bull., *2* (1959), 167–173.

Institute for Advanced Study
and
McGill University

SOME TWO-DIMENSIONAL UNITARY GROUPS GENERATED BY THREE REFLECTIONS

D. W. CROWE[1]

1. Introduction. Shephard and Todd **(5)** give generators for the finite primitive irreducible groups generated by two unitary reflections in $\mathbf{U}_2$. It is the purpose of the present paper to give generating reflections, and defining relations in terms of these reflections, for the seven such groups requiring three generating reflections, that is, for their nos. 7, 11, 12, 13, 15, 19, 22. The reflections are chosen whenever possible so that their product has the property suggested by Theorem 5.4 of **(5)**. That is, except for no. 15, the period of the product of the three generating reflections is $h = m_2 + 1$, and the characteristic roots of this product are $2\pi i m_1/h$ and $2\pi i m_2/h$, where m_1 and m_2 are the "exponents" **(5**, p. 282**)** of the group. The reason for the impossibility of such a choice for no. 15 is given in § 4. In § 5 the homomorphisms between these groups and certain groups of motions in elliptic 3-space are determined.

As in **(5)**, $\omega = \exp 2\pi i/3$, $\epsilon = \exp 2\pi i/8$, and $\eta = \exp 2\pi i/5$. The order of group $\mathfrak{G}$ is $|\mathfrak{G}|$. The identity element of a group, and the 2×2 identity matrix, are both designated by E. The notation $Z \rightleftarrows S, T$ means $ZS = SZ$ and $ZT = TZ$.

2. Groups 7, 11, 19. In terms of the generators

$$S = \omega\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}, \quad T = \frac{\omega\epsilon^5}{\sqrt{2}}\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}, \quad \text{and} \quad Z = -i\omega\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

the defining relations for no. 7 **(5**, pp. 280–1**)** are

2.1 $$S^2 = Z^8,\ T^3 = E,\ (ST)^3 = Z^4,\quad Z^{12} = E,\ Z \rightleftarrows S,T.$$

We let

2.2 $$R_1 = SZ^2,\ R_2 = T,\ R_3 = (STZ^3)^{-1}.$$

Then it can be readily verified that R_1, R_2, and R_3 are reflections, and that 2.1 and 2.2 imply

2.3 $$R_1^2 = R_2^3 = R_3^3 = E,\quad (R_1R_3)^2 = (R_3R_1)^2,$$

and

2.4 $$R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2,\quad T = R_2,\ S = (R_1R_2R_3)^2R_1,\ Z = (R_3R_1R_2)^{-1}.$$

Received January 4, 1960. This paper is a portion of the author's Ph.D. thesis at the University of Michigan, prepared under the direction of Professor H. S. M. Coxeter.

Conversely, 2.3 and 2.4 together imply 2.1 and 2.2. First, note that

$$Z^{12} = (R_3R_1R_2)^{-12} = (R_3R_1)^{-12}R_2^{-12} = E,$$

since 2.1 implies $(R_3R_1)^{12} = E$ (**2**, p. 77). Thus

$$S^2 = (R_1R_2R_3)^4R_1^2 = Z^{-4} = Z^8,$$

since $R_1 \rightleftarrows R_1R_2R_3$. Certainly

$$T^3 = R_2^3 = E, \quad \text{and} \quad (ST)^3 = [(R_1R_2R_3)^2R_1R_2]^3 = (R_3R_1R_2)^9 = Z^3,$$

since $R_3 \rightleftarrows R_1R_2$. Finally,

$$R_1 = SZ^2, \quad R_2 = T,$$

and

$$R_3 = R_2^{-1}R_1^{-1}Z^{-1} = T^{-1}S^{-1}Z^{-3} = (STZ^3)^{-1}.$$

In Table I we give the generating reflections and defining relations for the seven groups we consider. The proofs that the given relations for nos. 11 and 19 are equivalent to those given in (**5**) are so similar to the proof just given that they are omitted. In each case $Z = (R_1R_2R_3)^{-1}$. For no. 11, $(R_1R_3)^{24} = E$; for no. 19, $(R_1R_3)^{60} = E$.

If $\mathfrak{G} = \{R_1, R_3\}$ is an arbitrary group defined by relations between its two generators R_1 and R_3, and if the group $\mathfrak{H} = \{R_1, R_2, R_3\}$ is defined by the defining relations for $\mathfrak{G}$ together with $R_2^n = E$ and $R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2$ then $|\mathfrak{H}| = n|\mathfrak{G}|$. In the present case we can be more specific. Denote by $p_1[2n]p_2$ the group defined by

$$R_1^{p_1} = R_3^{p_2} = E,$$
$$R_1R_3R_1 \ldots = R_3R_1R_3 \ldots \text{ (}2n\text{ factors on each side) (\textbf{2}, p. 80).}$$

LEMMA. *If the period m of $(R_1R_3)^n$ in $p_1[2n]p_2$ is prime to n then the direct product $p_1[2n]p_2 \times \mathfrak{C}_n$ can be presented in the form*

$$R_1^{p_1} = R_2^n = R_3^{p_2} = E, (R_1R_3)^n = (R_3R_1)^n, R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2.$$

Proof. We need only show that we can find an element P in $\{R_1, R_2, R_3\}$ but not in $\{R_1, R_3\}$, such that P is of period n and $P \rightleftarrows R_1, R_3$. Since m is prime to n we can find some multiple of m, say r, such that $r \equiv 1 \bmod n$. Let $P = (R_3R_1)^rR_2$. Then $P^n = (R_3R_1)^{rn}R_2^n = E$, since $R_2 \rightleftarrows R_3R_1$. Using the fact that $R_1, R_3 \rightleftarrows (R_3R_1)^{r-1}$ we have $R_1P = R_1(R_3R_1)^rR_2 = (R_3R_1)^{r-1}R_1R_3R_1R_2 = (R_3R_1)^{r-1}R_1R_2R_3R_1 = PR_1$, and $R_3P = R_3(R_3R_1)^rR_2 = (R_3R_1)^{r-1}R_3R_3R_1R_2 = (R_3R_1)^{r-1}R_3R_1R_2R_3 = PR_3$. This completes the proof.

From this we get immediately

THEOREM 2.1.

$$\begin{array}{rl} \text{(i)} & \text{No. } 7 \cong 2[6]3 \times \mathfrak{C}_3 \\ \text{(ii)} & \text{No. } 11 \cong 2[6]4 \times \mathfrak{C}_3 \\ \text{(iii)} & \text{No. } 19 \cong 2[6]5 \times \mathfrak{C}_3. \end{array}$$

Proof. The values of n, m, r are as follows (**2**, p. 76):

$$\begin{array}{llll} \text{(i)} & n = 3, & m = r = 4 & \\ \text{(ii)} & n = 3, & m = 8, & r = 16 \\ \text{(iii)} & n = 3, & m = 20, & r = 40. \end{array}$$

3. Groups 12, 13, 22. The appropriate generators and defining relations for nos. 12, 13, and 22 appear in Table I. The given relations for no. 12 imply $(R_1R_3)^3 = (R_2R_3)^6 = E$. If the relation $(R_2R_3)^6 = E$ is replaced by $(R_2R_3)^3 = E$ the resulting group has half the order of no. 12. It is Coxeter's $[1\,1\,1]^2 \cong \mathfrak{S}_4$ (**1**, p. 248). The relations for no. 22 imply $(R_1R_3)^5 = E$. If the relation $(R_2R_3)^6 = E$ is replaced by $(R_2R_3)^3 = E$ the resulting group has half the order of no. 22. Slightly extending Coxeter's notation it is $[1\,1^5\,1^5]^2$. Although its order is 120 it is not $\mathfrak{S}_5$. By analogy with nos. 12 and 22 it might be expected that no. 13 could be defined by $S_i^2 = (S_1S_2)^4 = (S_1S_3)^4 = (S_2S_3)^6 = S_1S_2S_3S_2S_1S_3S_2S_3 = E$. However, there is no choice of three of the 18 reflections in no. 13 having products of these periods.

It can be verified directly (as was done for no. 7) that the tabulated relations for these three groups are equivalent to those of Shephard and Todd. However, these calculations are tedious and unenlightening. It is more convenient to use the method of enumeration of cosets (**2**, pp. 12–17). Enumeration of the 4 cosets of the subgroup $\{R_2, R_3\}$ (of order $\leqslant 12$) generated by R_2 and R_3 shows that the relations given for no. 12 define a group of order $\leqslant 4.12$. But since the generators S, T, Z are in this group the order is also $\geqslant 48$. Exactly similar arguments apply to nos. 13 and 22. In the former the subgroup $\{R_2, R_3\}$ is of order $\leqslant 16$ and has 6 cosets. In the latter the subgroup $\{R_2, R_3\}$ is of order $\leqslant 12$ and has 20 cosets.

An alternative set of generating reflections for no. 12 is $P_1 = R_1$, $P_2 = R_2$, $P_3 = R_1R_3R_1$. The corresponding defining relations are $P_i^2 = (P_1P_2)^3 = (P_1P_3)^3 = E$, $P_1(P_2P_3)^2 = (P_3P_2)^2P_1$. These imply $(P_2P_3)^4 = E$. Analogous generating reflections for no. 22 are $P_1 = -R_3R_1R_2R_1R_3$, $P_2 = -R_1$, $P_3 = R_2$. Defining relations are $P_i^2 = (P_1P_2)^3 = (P_1P_3)^3 = (P_2P_3)^{10} = E$, $P_1(P_2P_3)^2P_2 = (P_3P_2)^2P_3P_1$. These might be considered analogues of the tabulated definition for no. 13 since the relation $(R_1R_2)^2 = (R_2R_3)^4$ of no. 13 implies both $R_3(R_1R_2)^2 = (R_2R_1)^2R_3$ and $R_1(R_2R_3)^4 = (R_3R_2)^4R_1$.

4. Group 15. Generating reflections and defining relations for no. 15 appear in Table I. The sufficiency of the definition can be verified by enumerating the 6 cosets of $\{R_2, R_3\}$ (of order $\leqslant 48$).

The exponents of this group are 11 and 23. We proceed to show that no matrix in the group has characteristic roots $\exp 2\pi i\, 11/24$, $\exp 2\pi i\, 23/24$. Thus, *a fortiori*, no product of generating reflections has these characteristic roots. We first note that no. 12 is a subgroup of index 6 in no. 15. In fact no. 12 is generated by iS_1 and T_1; its only scalar matrices are $\pm E$. No. 15 is generated by iS_1 and $i\omega T_1$, and contains $Z = -i\omega E$. That is, the elements

of no. 15 are of the form MZ^n, $n = 0, 1, \ldots, 5$, where M is a matrix of no. 12. Now suppose $MZ^n = (\alpha_{rs})$ has characteristic roots $\pm \exp 2\pi i\, 11/24$ for some choice of M and n. This implies $\alpha_{11} + \alpha_{22} = m_{11} + m_{22} = 0$, where m_{11} and m_{22} are the diagonal entries of M. The 18 matrices in no. 12 having this property are its 12 reflections and

$$\pm\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \pm\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \pm\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The former have characteristic roots ± 1, the latter $\pm i$. Since no product of ± 1 or $\pm i$ by a power of $-i\omega = \exp 2\pi i/12$ yields $\pm \exp 2\pi i\, 11/24$, we conclude that no matrix in no. 15 has these characteristic roots.

5. Homomorphisms with Goursat's groups. In this section we assume knowledge of **(3)**. Groups $\mathfrak{L}$, $\mathfrak{R}$, $\mathfrak{l}$, $\mathfrak{r}$ of Clifford translations corresponding to each of the present groups can be determined from quaternion representations of R_1, R_2, R_3. These all appear in Table II. They determine groups 2:1 homomorphic to certain groups of motions in elliptic 3-space given by Goursat **(4)**. In fact, the latter groups are determined by isomorphisms $\mathfrak{L}'/\mathfrak{l}' \cong \mathfrak{R}'/\mathfrak{r}'$ where $\mathfrak{L}'$, $\mathfrak{R}'$, $\mathfrak{l}'$, $\mathfrak{r}'$ are the polyhedral or cyclic groups corresponding to the binary polyhedral or cyclic groups $\mathfrak{L}$, $\mathfrak{R}$, $\mathfrak{l}$, $\mathfrak{r}$ by 2:1 homomorphism.

The subgroups generated by pairs of generating reflections are groups of regular complex polygons. These have been found, after $\mathfrak{L}$ and $\mathfrak{R}$, by reference to the Table in **(3)**. There are some possible ambiguities in this determination, which can all be readily resolved. For example, the subgroup $\{R_2, R_3\}$ of no. 7 has $\mathfrak{L} \cong \mathfrak{C}_6$, $\mathfrak{R} \cong \langle 2, 3, 3\rangle$. Reference to the Table of **(3)** shows that this applies to either 3[4]3 or 3[3]3. But the generators

$$\omega i\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \frac{\omega\epsilon^5}{\sqrt{2}}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

(5, p. 281**)** of the larger group 3[4]3 are both in $\{R_2, R_3\}$. Therefore $\{R_2, R_3\}$ is 3[4]3.

We summarize the results of Table II. For a given group $\mathfrak{G}$ let the periods of the tabulated generating reflections R_1, R_2, R_3 be p_1, p_2, p_3. Let the collineation group of $\mathfrak{G}$ be $(2, 3, \nu)$. (For no. 7, $\nu = 3$; for nos. 11, 12, 13, 15, $\nu = 4$; for nos. 19, 22, $\nu = 5$.) Let $\mathfrak{Z}$ be the centre of $\mathfrak{G}$.

THEOREM 5.1. *The group $\mathfrak{G} = \{R_1, R_2, R_3\}$ is* 2:1 *homomorphic to the group of motions in elliptic* 3*-space defined by the isomorphism $\mathfrak{L}'/\mathfrak{l}' \cong \mathfrak{R}'/\mathfrak{r}'$ where*

(a) *$\mathfrak{L}'$ and $\mathfrak{l}'$ are cyclic groups.*

(b) $|\mathfrak{L}'| =$ l.c.m. $\{p_1, p_2, p_3\}$.

(c) $2|\mathfrak{l}'| = |\mathfrak{Z}|$. *That is, except for no.* 15, $2|\mathfrak{l}'|$ *is the period of the smallest power of $R_1R_2R_3$ which is central.*

(d) *$\mathfrak{R}'$ is $(2, 3, \nu)$.*

(e) *$\mathfrak{r}'$ is the unique normal subgroup of $\mathfrak{R}'$ such that $|\mathfrak{L}'|\,|\mathfrak{r}'| = |\mathfrak{R}'|\,|\mathfrak{l}'|$. In fact, except for no.* 12, *$\mathfrak{r}' \cong \mathfrak{R}'$ and $\mathfrak{l}' \cong \mathfrak{L}'$.*

TABLE I

GENERATING REFLECTIONS AND ABSTRACT DEFINITIONS FOR GROUPS GENERATED BY THREE REFLECTIONS IN U_2

Group	Generating reflections and notations of **(5)**	Defining relations	$R_1R_2R_3$
7	$R_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = SZ^2$	$R_1{}^2 = R_2{}^3 = R_3{}^3 = E$	
	$R_2 = \frac{\omega\epsilon^5}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} = T$	$(R_1R_3)^3 = (R_3R_1)^3$ $R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2$	$\exp 2\pi i \frac{11}{12} E$
	$R_3 = \frac{\omega\epsilon^3}{\sqrt{2}} \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} = (STZ^3)^{-1}$		
11	$R_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ i & -1 \end{pmatrix} = SZ^{12}$	$R_1{}^2 = R_2{}^3 = R_3{}^4 = E$	
	$R_2 = \frac{\omega\epsilon^5}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = TZ^9$	$(R_1R_3)^3 = (R_3R_1)^3$ $R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2$	$\exp 2\pi i \frac{23}{24} E$
	$R_3 = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} = (ST)^{-1}Z^2$		
12	$R_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & i \\ -1 & 1 \end{pmatrix} = S$		$\frac{i}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix}$ Char. roots:
	$R_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = T^3ST$	$R_i{}^2 = (R_1R_2)^3$ $= R_1R_2R_3R_2R_1R_3R_2R_3 = E$	$\exp 2\pi i \frac{5}{8}$
	$R_3 = \epsilon \begin{pmatrix} 0 & -i \\ 1 & 0 \end{pmatrix} = TST^2$		$\exp 2\pi i \frac{7}{8}$
13	$R_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = T^3STZ$	$R_i{}^2 = (R_1R_2)^3 = (R_1R_3)^4 = E$	$\frac{\epsilon^7}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ Char. roots:
	$R_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & i \\ -i & 1 \end{pmatrix} = S$	$(R_1R_3)^2 = (R_2R_3)^4$	$\exp 2\pi i \frac{7}{12}$
	$R_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = (TS)^2Z^2$		$\exp 2\pi i \frac{11}{12}$

TABLE 1 (*Cont.*)

Group	Generating reflections and notations of **(5)**	Defining relations	$R_1R_2R_3$
15	$R_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & i \\ -i & 1 \end{pmatrix} = S$	$R_1^2 = R_2^2 = R_3^3 = E$	
	$R_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = (ST)^2Z^7$	$(R_2R_3)^3 = (R_3R_2)^3$ $R_2R_1R_3 = R_1R_3R_2$	
	$R_3 = \frac{\omega\epsilon^5}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = TZ^3$	$R_3R_1R_3R_1R_2 = R_1R_2R_3R_1R_3$	
19	$R_1 = \frac{i}{\sqrt{5}}\begin{pmatrix} \eta - \eta^4 & \eta^3 - \eta^2 \\ \eta^3 - \eta^2 & \eta^4 - \eta \end{pmatrix} = SZ^{10}$	$R_1^2 = R_2^3 = R_3^5 = E$	
	$R_2 = \frac{\omega}{\sqrt{5}}\begin{pmatrix} \eta^2 - \eta^4 & \eta^4 - 1 \\ 1 - \eta & \eta^3 - \eta \end{pmatrix} = TZ^{29}$	$(R_1R_3)^3 = (R_3R_1)^3$	$\exp 2\pi i \frac{59}{60} E$
	$R_3 = \begin{pmatrix} \eta^4 & 0 \\ 0 & 1 \end{pmatrix} = (ST)^{-1}Z^{20}$	$R_1R_2R_3 = R_2R_3R_1 = R_3R_1R_2$	
22	$R_1 = \frac{i}{\sqrt{5}}\begin{pmatrix} \eta^4 - \eta & \eta^2 - \eta^3 \\ \eta^2 - \eta^3 & \eta - \eta^4 \end{pmatrix} = S$		$\frac{i}{\sqrt{5}}\begin{pmatrix} \eta^4 - 1 & 1 - \eta^3 \\ \eta^2 - 1 & \eta - 1 \end{pmatrix}$ Char. roots:
	$R_2 = \frac{i}{\sqrt{5}}\begin{pmatrix} \eta - \eta^4 & \eta^4 - \eta^3 \\ \eta^2 - \eta & \eta^4 - \eta \end{pmatrix} = STST^2SZ^2$	$R_i^2 = (R_1R_2)^5 = (R_2R_3)^6$ $= R_1R_2R_3R_2R_1R_3R_2R_3 = E$	$\exp 2\pi i \frac{11}{20}$
	$R_3 = \frac{i}{\sqrt{5}}\begin{pmatrix} \eta - \eta^4 & \eta^2 - \eta \\ \eta^4 - \eta^3 & \eta^4 - \eta \end{pmatrix} = T^3STZ^2$		$\exp 2\pi i \frac{19}{20}$

TABLE II*

GOURSAT GROUPS CORRESPONDING TO GROUPS GENERATED BY THREE REFLECTIONS IN U_2

Group		b	$\frac{\mathfrak{L}}{\mathfrak{l}}$	$\frac{\mathfrak{R}}{\mathfrak{r}}$	Subgroups $\{R_1, R_2\}$, $\{R_1, R_3\}$, $\{R_2 R_3\}$, Subgroup	$\mathfrak{L}$	$\mathfrak{R}$	$p_1[q]p_2$
7	R_1:	$-\mathbf{i}$			$\{R_1, R_2\}$	$\mathfrak{C}_{12}$	$\langle 2,3,3\rangle$	3[6]2
	R_2:	$\frac{1}{2}+\frac{\mathbf{i}}{2}-\frac{\mathbf{j}}{2}+\frac{\mathbf{k}}{2}$	$\frac{\mathfrak{C}_{12}}{\mathfrak{C}_{12}}$	$\frac{\langle 2,3,3\rangle}{\langle 2,3,3\rangle}$	$\{R_1, R_3\}$	$\mathfrak{C}_{12}$	$\langle 2,3,3\rangle$	3[6]2
	R_3:	$\frac{1}{2}+\frac{\mathbf{i}}{2}-\frac{\mathbf{j}}{2}-\frac{\mathbf{k}}{2}$			$\{R_2, R_3\}$	$\mathfrak{C}_6$	$\langle 2,3,3\rangle$	3[4]3
11	R_1:	$\frac{\mathbf{i}}{\sqrt{2}}-\frac{\mathbf{j}}{\sqrt{2}}$			$\{R_1, R_2\}$	$\mathfrak{C}_{12}$	$\langle 2,3,4\rangle$	3[8]2
	R_2:	$\frac{1}{2}+\frac{\mathbf{i}}{2}+\frac{\mathbf{j}}{2}+\frac{\mathbf{k}}{2}$	$\frac{\mathfrak{C}_{24}}{\mathfrak{C}_{24}}$	$\frac{\langle 2,3,4\rangle}{\langle 2,3,4\rangle}$	$\{R_1, R_3\}$	C_8	$\langle 2,3,4\rangle$	4[6]2
	R_3:	$\frac{1}{\sqrt{2}}+\frac{\mathbf{i}}{\sqrt{2}}$			$\{R_2, R_3\}$	$\mathfrak{C}_{24}$	$\langle 2,3,4\rangle$	4[4]3
12	R_1:	$\frac{\mathbf{i}}{\sqrt{2}}+\frac{\mathbf{j}}{\sqrt{2}}$			$\{R_1, R_2\}$	$\mathfrak{C}_4$	$\langle 2,2,3\rangle$	2[3]2
	R_2:	$\frac{-\mathbf{i}}{\sqrt{2}}-\frac{\mathbf{k}}{\sqrt{2}}$	$\frac{\mathfrak{C}_4}{\mathfrak{C}_2}$	$\frac{\langle 2,3,4\rangle}{\langle 2,3,3\rangle}$	$\{R_1, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,3\rangle$	2[3]2
	R_3:	$\frac{-\mathbf{j}}{\sqrt{2}}-\frac{\mathbf{k}}{\sqrt{2}}$			$\{R_2, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,3\rangle$	2[6]2
13	R_1:	$\frac{-\mathbf{i}}{\sqrt{2}}-\frac{\mathbf{k}}{\sqrt{2}}$			$\{R_1, R_2\}$	$\mathfrak{C}_4$	$\langle 2,2,3\rangle$	2[3]2
	R_2:	$\frac{\mathbf{i}}{\sqrt{2}}+\frac{\mathbf{j}}{\sqrt{2}}$	$\frac{\mathfrak{C}_4}{\mathfrak{C}_4}$	$\frac{\langle 2,3,4\rangle}{\langle 2,3,4\rangle}$	$\{R_1, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,2\rangle$	2[4]2
	R_3:	$\mathbf{j}$			$\{R_2, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,4\rangle$	2[8]2

*When a reflection R has period n, its quaternion form is $q' = aqb$, where $a = \exp 2\pi\mathbf{i}/n$ and b is given in the second column below.

TABLE II* (*Cont.*)

*When a reflection R has period n, its quaternion form is $q' = aqb$, where $a = \exp 2\pi\mathbf{i}/n$ and b is given in the second column below.

TABLE II* (*Cont.*)

Group		b	$\mathfrak{L}/\mathfrak{l}$	$\mathfrak{R}/\mathfrak{r}$	Subgroups $\{R_1, R_2\}$, $\{R_1, R_3\}$, $\{R_2, R_3\}$: Subgroup	$\mathfrak{L}$	$\mathfrak{R}$	$p_1[q]p_2$
15	R_1:	$\frac{\mathbf{i}}{\sqrt{2}} + \frac{\mathbf{j}}{\sqrt{2}}$			$\{R_1, R_2\}$	$\mathfrak{C}_4$	$\langle 2,2,4\rangle$	2[8]2
	R_2:	$-\mathbf{i}$	$\frac{\mathfrak{C}_{12}}{\mathfrak{C}_{12}}$	$\frac{\langle 2,3,4\rangle}{\langle 2,3,4\rangle}$	$\{R_1, R_3\}$	$\mathfrak{C}_{12}$	$\langle 2,3,4\rangle$	3[8]2
	R_3:	$\frac{1}{2} + \frac{\mathbf{i}}{2} + \frac{\mathbf{j}}{2} + \frac{\mathbf{k}}{2}$			$\{R_2, R_3\}$	$\mathfrak{C}_{12}$	$\langle 2,3,3\rangle$	3[6]2
19†	R_1:	$\mathbf{i}\frac{(2+\tau)^{\frac{1}{2}}}{\sqrt{5}} - \mathbf{k}\frac{(3-\tau)^{\frac{1}{2}}}{\sqrt{5}}$			$\{R_1, R_2\}$	$\mathfrak{C}_{12}$	$\langle 2,3,5\rangle$	3[10]2
	R_2:	$\frac{1}{2} + \mathbf{i}\frac{(4\tau+3)^{\frac{1}{2}}}{2\sqrt{5}}$			$\{R_1, R_3\}$	$\mathfrak{C}_{20}$	$\langle 2,3,5\rangle$	5[6]2
		$+ \mathbf{j}\frac{3-\tau}{2\sqrt{5}} + \mathbf{k}\frac{(2+\tau)^{\frac{1}{2}}}{2\sqrt{5}}$	$\frac{\mathfrak{C}_{60}}{\mathfrak{C}_{60}}$	$\frac{\langle 2,3,5\rangle}{\langle 2,3,5\rangle}$	$\{R_2, R_3\}$	$\mathfrak{C}_{30}$	$\langle 2,3,5\rangle$	5[4]3
	R_3:	$-\eta^2$						
22†	R_1:	$-\mathbf{i}\frac{(2+\tau)^{\frac{1}{2}}}{\sqrt{5}} + \mathbf{k}\frac{(3-\tau)^{\frac{1}{2}}}{\sqrt{5}}$			$\{R_1, R_2\}$	$\mathfrak{C}_4$	$\langle 2,2,5\rangle$	2[5]2
	R_2:	$\mathbf{i}\frac{(2+\tau)^{\frac{1}{2}}}{\sqrt{5}} + \frac{\mathbf{j}}{2} - \mathbf{k}\frac{(7-4\tau)^{\frac{1}{2}}}{2\sqrt{5}}$	$\frac{\mathfrak{C}_4}{\mathfrak{C}_4}$	$\frac{\langle 2,3,5\rangle}{\langle 2,3,5\rangle}$	$\{R_1, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,5\rangle$	2[5]2
	R_3:	$\mathbf{i}\frac{(2+\tau)^{\frac{1}{2}}}{\sqrt{5}} - \frac{\mathbf{j}}{2} - \mathbf{k}\frac{(7-4\tau)^{\frac{1}{2}}}{2\sqrt{5}}$			$\{R_2, R_3\}$	$\mathfrak{C}_4$	$\langle 2,2,3\rangle$	2[6]2

*When a reflection R has period n, its quaternion form is $q' = aqb$, where $a = \exp 2\pi\mathbf{i}/n$ and b is given in the second column below.
†Here $\tau = (1 + \sqrt{5})/2$.

REFERENCES

1. H. S. M. Coxeter, *Groups generated by unitary reflections of period two*, Can. J. Math., *9* (1957), 243–272.
2. H. S. M. Coxeter and W. O. J. Moser, *Generators and reflections for discrete groups* (Berlin, 1957).
3. D. W. Crowe, *The groups of regular complex polygons*, Can. J. Math., *13* (1961), 149–156.
4. E. Goursat, *Sur les substitutions orthogonales et les divisions régulières de l'espace*, Ann. Sci. Ecole Norm. Sup. (3), *6* (1889), 9–102.
5. G. C. Shephard and J. A. Todd, *Finite unitary reflection groups*, Can. J. Math., *6* (1954), 274–304.

University College, Ibadan

MOULTON PLANES

WILLIAM A. PIERCE

1. Introduction.[1] In 1902, F. R. Moulton **(12)** gave an early example of a non-Desarguesian plane. Its "points" are ordered pairs (x, y) of real numbers. Its "lines" coincide with lines of the real affine plane *except that* lines of negative slope are "bent" on the x-axis, line $\{y = b + mx\}$, for negative m, being replaced by $\{y = b + mx \text{ if } y \leq 0, y = [m/2]\cdot[x + (b/m)] \text{ if } y > 0\}$. A certain Desarguesian configuration in the classical plane is shifted just enough to vitiate Desargues' Theorem for Moulton's geometry. The plane is neither a translation plane ("Veblen–Wedderburn" in the sense of Hall **(7)**, p. 364) nor even the dual of one (Veblen and Wedderburn **(17)**. It is natural to ask if the same construction is feasible when real numbers are replaced by elements from an arbitrary field. If the construction *does* work—what geometric properties, what co-ordinate systems, and what collineation groups are obtained? Are the planes essentially "new"? In this paper and in a forthcoming sequel, I construct "Moulton planes" over a wide class of fields and answer relevant questions about their geometries. The classical ordering is replaced by a generalization of positives and negatives—the appropriate concept being that of "pseudo-order" used earlier by Dickson **(5)**, Kustaanheimo **(9)** and **(10)**, Pickert **(13)**, Sperner **(16)**, and others. The "positive" elements of F shall consist of a multiplicative subgroup P having index 2; and the "negatives," the other coset of non-zero elements. The product of two "positives," or of two "negatives," is still "positive"—while the product of a "negative" and a "positive" is "negative." (Write $x > 0$ or $x < 0$ according as $x \in P$ or $x \notin P \cup \{0\}$; say that non-zero elements x and y have the same or opposite "sign" according as $x/y > 0$ or < 0.) The field F is ordered in the usual sense if and only if P is closed under addition. The "pseudo-ordered" fields include ordered fields as special cases: rationals, reals, etc., under the standard order. A single field, F, may admit more than one "pseudo-order." For example, an unfamiliar definition of "positive" and "negative" exists on the rationals as follows. Given any (rational) prime p, a rational number, r, is uniquely expressible in the form $(p^i a)/b$ where i is integral, a and b denote (rational) integers prime to p, and a/b is reduced to lowest terms. For $a \neq 0$, one may call r "positive" or "negative" depending on whether i is even or odd. The rationals are *not ordered* under this definition of "pseudo-order." (For instance—given a prime p, and any non-zero integer

Received August 24, 1959.

[1]This is to express my gratitude to Professor Günter Pickert for helpful suggestions, simplifications, and extensions—especially the connection between this paper and cartesian groups.

b prime to p, $[(p-1)/b] + [1/b] = p/b$, showing that the sum of two "positives" can be "negative.")

Another non-trivial "pseudo-order" can be constructed as follows. Let $F(x)$ denote the field of rational forms over a field F. A quotient $f(x)/q(x)$, reduced to lowest terms with $f(x)\cdot q(x) \neq 0$, is > 0 or < 0 according as the difference of degrees, $\delta(f) - \delta(q)$, is even or odd.

A multiplicative subgroup P of index 2 must contain *all* non-zero squares. On the other hand, P may or may not consist *only* of squares. Under the usual ordering, positive real numbers are all squares—positive rationals are not. (Note that the "positive" rationals are not necessarily squares under the alternative "pseudo-order" described above.) What if F is finite? In the case of characteristic 2, $x \to x^2$ is an automorphism, F contains *only squares*, and *no* "pseudo-order" is possible. In finite fields of odd characteristic, however, the powers of a primitive element indicate that *at least half* of the non-zero elements are squares; since any equation $x^2 = a_0^2$, with $a_0 \neq 0$, has two distinct solutions, the map $x \to x^2$ cannot be "onto"; so the squares form a *proper* multiplicative subgroup—and that subgroup has index 2.

Dickson **(5)**, studying equations over finite fields, used non-zero squares as "positives" in his effective treatment of discriminants. Sperner **(16)** used the more general concept of "pseudo-order" to investigate relations between algebraic semiorder and geometric order. Recently, Kustaanheimo **(9)** has utilized the same concept to develop order and congruence relations for finite geometries. He has suggested the intriguing possibility that such ideas may be applied to problems of quantum physics—where some of the difficulties encountered are not necessarily intrinsic, but may stem from the imposition of infinite models on finite situations.

My own interest in Moulton's construction—especially over finite fields—has motivated Carlitz to prove two basic theorems. The statements of his results require a preliminary definition, which will also be needed later.

Definition 1. A single-valued function ϕ on a field F is called "order-preserving" if and only if $[\phi(u) - \phi(v)]/(u-v) > 0$, for all distinct $u, v \in F$; "monotonic" if $[\phi(u) - \phi(v)]/(u-v)$ retains the same "sign," for all $u \neq v \in F$.

(A). (Carlitz **(3)**). If F denotes a finite field of odd characteristic, the most general one-to-one monotonic function, ϕ, on F is given by $\phi(x) = a\cdot\sigma(x) + b$ (where σ is an automorphism; $b, a \in F$ with $a \neq 0$). According as $a > 0$ or $a < 0$, ϕ is order-preserving or reversing. If $\phi(0) = 0$, $\phi(1) = 1$, then ϕ is an automorphism.

(B). (A generalization of the above, unpublished, but contained in a written communication to me.) Assume that F has order p^n, where p is odd. Put

$$\psi(a) = a^{\frac{1}{2}(p^n-1)},$$

and let $\lambda_1 = \pm 1, \ldots, \lambda_k = \pm 1$. Let $f(x_1, \ldots, x_k)$ be a polynomial with

coefficients in $GF(p^n)$ such that $\psi\{f(x_1, \ldots, x_r, \ldots, x_k) - f(x_1, \ldots, y_r, \ldots, x_k)\} = \lambda_r\psi(x_r - y_r)$ for all $r = 1, \ldots, k$ and all x_j, y_j in $GF(p^n)$. Then

$$f(x_1, \ldots, x_k) = c_1x_1^{p^{r_1}} + \ldots + c_kx_k^{p^{r_k}} + d,$$

where $\psi(c_j) = \lambda_j$ and $0 \leqslant r_j < n$.

2. Definition and construction of "Moulton planes." Throughout this paper, I shall assume that a "pseudo-order"—hence a multiplicative subgroup P of index 2—exists and has been specified on a given field F. Terms and symbols of "order," "inequality," etc. will refer to the designated "pseudo-order"; they will no longer be enclosed by quotation marks. It will be convenient to replace the x-axis by the y-axis as the line along which "bending" occurs.

Definition 2. Let ϕ denote a one-to-one function of a given field F onto itself. A *Moulton construction*, $C_\phi(F)$, consists of "points" and classes of "points"—called "lines"—in which:

(i) Each "point" is an ordered pair (x, y) of elements $\in F$.

(ii) Each "line" consists of all "points" (x, y) satisfying an equation of the form $\{x = c\}(c \in F)$, or $\{y = b + m \circ x\}(b, m \in F)$, where $m \circ x$ is defined from the field multiplication by $m \circ x = mx$ or $\phi(m) \cdot x$ according as $x \geqslant 0$ or $x < 0$ [a "line" of the latter type is said to have "slope" m].

Definition 3. A *Moulton plane* is a construction, $C_\phi(F)$, whose "points" and "lines" form an affine plane. If $C_\phi(F)$ is such a plane, it will be denoted by $M_\phi(F)$.

Remark. When convenient, $M_\phi(F)$ will also be regarded as the projective plane obtained by adding ideal elements to the affine Moulton plane.

3. Geometry of Moulton planes.

THEOREM 1. *A construction, $C_\phi(F)$ forms a Moulton plane if and only if:*

(a) *The function ϕ is order-preserving.*

(b) *Given any negative $n_0 \in F$, $x \to [\phi(x) - n_0x]$ maps F onto F.*

Proof. Two distinct "points" determine a unique "line" except possibly in the case of non-zero abscissae having unlike signs. Given $u_0 < 0$, $p_0 > 0$, $v_0, q_0 \in F$, the existence of *at least* one "line" $(u_0, v_0) \cup (p_0, q_0)$ amounts to the existence of $m \in F$ such that (u_0, v_0) satisfies $y = \phi(m) \cdot x + (q_0 - mp_0)$, that is, of an m for which $\phi(m) - [p_0/u_0] \cdot m = (v_0 - q_0)/u_0$. Such an m exists for all (u_0, v_0) and (p_0, q_0), with $u_0 < 0$, $p_0 > 0$, if and only if condition (b) holds. Suppose (Fig. 1) that both (u_0, v_0) and (p_0, q_0) belong to lines of "slope" m and n, whence $v_0 = \phi(m) \cdot u_0 + (q_0 - mp_0)$ and

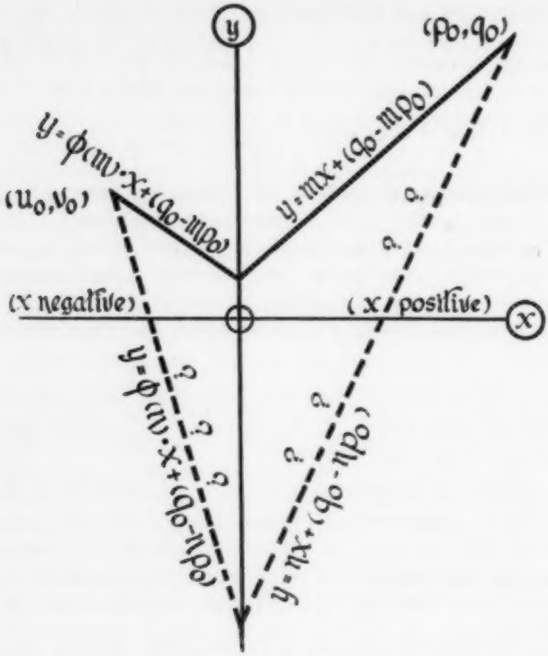

FIG. 1.

$v_0 = \phi(n) \cdot u_0 + (q_0 - np_0)$. Subtraction gives $0 = u_0 \cdot [\phi(m) - \phi(n)] - p_0 \cdot [m - n]$. Unless $m - n = 0$, we get $[\phi(m) - \phi(n)]/[m - n] = p_0/u_0 < 0$. Thus, order-preservation is sufficient to prove that *not more than* one "line" joins any two distinct "points." Conversely, the existence of $m \neq n$ such that $[\phi(m) - \phi(n)]/[m - n] = 1/r_0 < 0$ would permit us to put $v_0 = \phi(m) \cdot r_0 - m = \phi(n) \cdot r_0 - n$, forcing (r_0, v_0) to lie on distinct "lines" of "slope" m and n through $(1, 0)$. It follows that order-preservation is equivalent to the existence of *at most* one "line" through any two distinct "points."

Let us now verify Euclid's parallel postulate. Two "lines" are "parallel" if and only if they coincide or have no point in common. An ordinary "point" (x_0, y_0) must be shown to lie on exactly one "line" parallel to an ordinary "line" l. (i) If l is given by $\{x = c\}$, (x_0, y_0) lies on $\{x = a\}$ if and only if $a = x_0$. On the other hand $\{x = c\}$ intersects *every* "line" of the form $\{y = b + m \circ x\}$. (ii) If l is given by $\{y = b + m \circ x\}$, then (x_0, y_0) lies on $\{y = c + m \circ x\}$ if and only if $c = y_0 - m \circ x_0$. For $m \neq n$, $\{y = b + m \circ x\}$ meets $\{y = d + n \circ x\}$ in the point $(u_0, b + m \circ u_0)$, with $u_0 = (b - d)/(n - m)$ or $(b - d)/[\phi(n) - \phi(m)]$ according as $u_0 \geqslant 0$ or $u_0 < 0$. Such a u_0 exists

since, according to (a), $(b - d)/(n - m)$ and $(b - d)/[\phi(n) - \phi(m)]$ have the same sign.

The presence of three non-collinear "points" is trivial—(0, 0), (1, 0), (0, 1) for example.

COROLLARY 1. *If F is finite (of odd characteristic), and if ϕ is one-to-one on F, then $C_\phi(F)$ is a plane if and only if ϕ preserves order.*

Proof. If ϕ fails to preserve order, the theorem shows that $C_\phi(F)$ cannot be a plane.

Assume, conversely, that ϕ *does* preserve order. Given

$$n_0 < 0, x \to [\phi(x) - n_0 x]$$

is one-to-one "into":

$$\phi(u) - n_0 u = \phi(v) - n_0 v \to \phi(u) - \phi(v) = n_0 \cdot (u - v),$$

which is impossible unless $u = v$. By finiteness, one-to-one "into" is "onto."

COROLLARY 2. *If F is a finite field (of odd characteristic), and if ϕ is a one-to-one function on F, then $C_\phi(F)$ is a Moulton plane if and only if $\phi(m) = a^2 \cdot \sigma(m) + b$, for some $b, a \neq 0 \in F$, and some automorphism σ. In case $\phi(0) = 0$ and $\phi(1) = 1$, a plane is obtained if and only if ϕ is an automorphism.*

Proof. This is a restatement of Corollary 1 in the presence of the Carlitz Theorem **(3)**.

Examples over the real field R (relative to the usual order). A construction $C_\phi(R)$ is a Moulton plane if and only if ϕ is an increasing function of R onto itself—for instance:

(1) $\phi(m) = m^3$,

(2) $\phi(m) = m$ or $p_0 m (p_0 > 0)$ according as $m \geqslant 0$ or $m < 0$ (the example originally given by Moulton); see also Pickert **(13)**, p. 93, *et seq.*

(3)
$$\phi(m) = \begin{cases} -2 - \sqrt{-m}, \text{ for } m < 0 \\ -2 + m, \text{ for } 0 \leqslant m < 1 \\ - - - - - - - \\ [(-r^2 - r - 4)/2] + (r + 1) \cdot m, \text{ for } r \leqslant m < r + 1, \\ \text{where } r \text{ is a non-negative integer.} \end{cases}$$

LEMMA 1. *Any Moulton plane, $M_\phi(F)$, is isomorphic to a plane $M_{(\phi')}(F)$ with $\phi'(0) = 0$, $\phi'(1) = 1$.*

Proof. Initially, "lines" are given by $\{x = c\}$ for $c \in F$, and $\{y = b + mx\}$ if $x \geqslant 0$; $\{y = b + \phi(m) \cdot x\}$ if $x < 0$. Change co-ordinates, putting $x = x'$, $y = y'$, if $x \geqslant 0$; and $x = cx'$, $y = y' + ax'$, if $x > 0$, where $a = \phi(0)/[\phi(1) - \phi(0)]$ and $c = 1/[\phi(1) - \phi(0)]$ with $c > 0$ since ϕ is order-preserving. This transformation permutes "lines" $\{x = c\}$ among themselves, maps $\{y = b + mx\}$

onto $\{y' = b + mx'\}$ for $x \geqq 0$, and $\{y = b + \phi(m) \cdot x\}$ onto $\{y' + ax' = b + c \cdot \phi(m) \cdot x'\}$ for $x < 0$; the latter reducing to $\{y' = b + [\phi'(m)] \cdot x'\}$, where $\phi'(m) = c \cdot \phi(m) - a$, $\phi'(0) = 0$, and $\phi'(1) = 1$.

THEOREM 2. *Every Moulton plane can be represented by co-ordinates from a Cartesian group, G (in the sense of Pickert* **(13)**, p. 90). *Addition for G coincides with that of F, but the multiplication,* $\circ$, *of G is defined as follows:* $u \circ v = uv$ or $\phi(u) \cdot v$, *according as* $v > 0$ *or* $v < 0$.

Proof. Apply Lemma 1 to represent the given plane as $M_\phi(F)$, where $\phi(0) = 0$, $\phi(1) = 1$. Since the elements of F already form a group under $+$, they will form a Cartesian group under the operations $\{+, \circ\}$ if and only if:

(i) The non-zero elements form a loop under $\circ$.
(ii) $x \in F \rightarrow 0 \circ x = x \circ 0 = 0$, $1 \circ x = x \circ 1 = x$.
(iii) For all $a, b, c, d \in F$, $a \circ c - a \circ d = b \circ c - b \circ d \rightarrow a = b$ or $c = d$ [Pickert **(13)**, p. 90 (9)].
(iv) Given $a, b, c \in F$, with $a \neq b$, $\exists\, x \in F$ such that $a \circ x - b \circ x = c$ **(13)**, p. 90 (10)].
(v) Given $a, b, c \in F$, with $a \neq b$, $\exists\, x \in F$ such that $x \circ a - x \circ b = c$. [**(13)**, p. 90 (11)].

In the presence of $\phi(0) = 0$ and $\phi(1) = 1$, properties (i) and (ii) are trivial. Property (iii) is immediate if $c/d > 0$, because $a \cdot (c - d) = b \cdot (c - d)$ or $\phi(a) \cdot (c-d) = \phi(b) \cdot (c-d)$ according as $c>0$ or $c<0$. To prove (iii) when $c/d < 0$, use the symmetry between c and d to suppose $c < 0$, $d > 0$. Then $\phi(a) \cdot c - ad = \phi(b) \cdot c - bd \rightarrow [\phi(a) - \phi(b)] \cdot c = (a - b) \cdot d$. Unless $a = b$, $[\phi(a) - \phi(b)]/(a - b) = d/c < 0$, contradicting the order-preservation.

To verify (iv), use $x = c/(a - b)$ if $c/(a - b) > 0$; otherwise, $x = c/[\phi(a) - \phi(b)]$.

Property (v) is obvious if $ab > 0$. Otherwise, after possible multiplication by $n < 0$, we can assume $a < 0$, $b > 0$. Property (b) of Theorem 1 asserts that the map $x \rightarrow \phi(x) - (b/a) \cdot x$ is "onto," thus supplying the desired value of x. The representation of lines follows at once from the Moulton construction, and the proof of the Theorem is complete.

Remark. The basic geometry of Moulton planes can be developed using direct, synthetic proofs. It is more efficient, however, to apply known results concerning Cartesian groups, as given by Pickert in *Projektive Ebenen* **(13)**. Identify Moulton points $(0, 0)$, $(1, 1)$, X_∞ (the ideal point on the x-axis), Y_∞ (the ideal point on the y-axis), and the infinite point on $\{y = x\}$, with the respective points O, E, U, V, W, of Pickert's co-ordinate system, and the (Moulton) ideal line, l_∞, with line $U \cup V$ (**(13)**, pp. 31–32). Put the Hall ternary (**(13)**, p. 35; Hall **(6)**), $T(u, x, v) = (u \circ x) + v$, so that Moulton lines have equations of the form $\{x = c\}$ and $\{y = T(m, x, b)\}$ (Pickert **(13)**, p. 35).

THEOREM 3. *Every Moulton plane ,M, is a Baer plane* (Baer **(1)**), *in the sense that it satisfies the Desarguesian* (Y_∞, l_∞)*-Theorem* (*Pickert* **(13)**, *pp.* 74–76). *Thus, M also satisfies the Reidemeister-condition for the* (X_∞, Y_∞, W)*-web* ("Gewebe")—(**(13)**, *p.* 52; *Reidemeister* **(15)**).

Proof. By Theorem 2, M can be co-ordinated over a Cartesian group. The present Theorem is then a restatement of Pickert's *Satz 36* (**(13)**, top of p. 100).

Note. The direct proof of Theorem 3 would present a neat geometric picture —the y-axis being used as an auxiliary line.

THEOREM 4. *In a Moulton plane* $M_\phi(F)$, *with* $\phi(0) = 0$, $\phi(1) = 1$, *the following assertions are equivalent:*

(a) *The Desarguesian* $(X_\infty, l_\infty; Y_\infty, \{y = 0\})$*—Theorem holds* [this involves triangles perspective from X_∞, with one pair of corresponding vertices, say P and P', on $\{y = 0\}$; a pair of corresponding sides, QR and $Q'R'$, through Y_∞; and PR parallel to $P'R'$ if and only if PQ is parallel to $P'Q'$].

(b) *The plane* $M_\phi(F)$ *is a translation plane with axis* l_∞ (*Pickert* **(13)**, p. 199; *Hall* **(7)**, p. 364–*a "Veblen–Wedderburn" plane in the latter's terminology*).

(c) *Desargues' Theorem is valid.*

(d) *The Cartesian group* $\{+, \circ\}$ *satisfies the right-distributive law* $u \circ (x + w) = (u \circ x) + (u \circ w)$, [**(13)**, p. 99, (18)].

(e) *The function* ϕ *is the Identity*! (Cf. footnote[2].)

Proof. Much of this Theorem is an immediate consequence of *Satz 37* (**(13)**, p. 100). By Theorem 3, $M_\phi(F)$ satisfies the Reidemeister-condition relative to $(U = X_\infty, V = Y_\infty, W)$—whence Pickert's condition (b) of *Satz 37* reduces to condition (a) above. By the associativity of addition, and by the "erste Zerlegbarkeitsbedingung," $T(u, x, v) = u \circ x + v$, condition (c) of *Satz 37* reduces to (d) of the present Theorem. (Cf. Satz, 35, p. 99.) Each of (a) and (d) becomes equivalent to the Desarguesian (Q, l_∞)-Theorem for two distinct choices—in this case X_∞ and Y_∞—of $Q \in l_\infty$, implying condition (b) of the present Theorem.

It remains only to show that (d) → (e), since the Theorem will then follow trivially. Suppose $\phi \neq \mathcal{J}$ (the identity), and let $\phi(u) \neq u$, for $u \in F$. If $x < 0$, we get $(u \circ x) + (u \circ 1) = \phi(u) \cdot x + u$, and $u \circ (x + 1) = \phi(u) \cdot (x + 1)$ or $u \cdot (x + 1)$, neither of which equals $[\phi(u) \cdot x + u]$. Thus, $\phi \neq \mathcal{J}$ implies that the right-distributive law cannot hold, and (e) follows from (d).

Note. A direct verification of Theorem 4 could be based on the following neat proof that (a) → (e). Suppose ϕ non-trivial. Choose u and n such that $\phi(u) \neq u$, $n < 0$, and $n + 1 \leqq 0$. Consider the triangles with vertices, $(1, 0)$, $(0, 1)$, $(0, b)$; and $(n + 1, 0)$, $(n, 1)$, $(n, -u)$. As triangles in the classical

[2]The redundance of Theorem 4 may help to clarify the relation between this development and that of Pickert **(13)**. Henceforth, all references will be to the latter work unless otherwise specified.

plane over F, they are perspective from X_∞, axial from l_∞, have a pair of corresponding vertices on $\{y = 0\}$ and a pair of corresponding sides through Y_∞. In $M_\phi(F)$, all these properties still hold *except that* $(1, 0) \cup (0, -u)$ and $(n + 1, 0) \cup (n, -u)$ have respective Moulton "slopes" u and $\phi^{-1}(u)$—violating the $(X_\infty, l_\infty; Y_\infty, \{y = 0\})$-condition of (a).

THEOREM 5. *Let $M = M_\phi(F)$ denote a Moulton plane where $\phi(0) = 0$, $\phi(1) = 1$. Each of the following is necessary and sufficient for M to be (Y_∞, Y_∞)-transitive:*

(i) *The (Y_∞, n)-Desargues' Theorem holds for every line n through Y_∞.*
(ii) *The left-distributive law $(a + b) \circ c = a \circ c + b \circ c$ is valid.*
(iii) *The Cartesian group $\{+, \circ\}$ is a left quasi-field.*
(iv) *The function ϕ is additive.*

Proof. Condition (i) is a standard variation of (Y_∞, Y_∞)-transitivity. Conditions (ii) and (iii) involve *Satz 39* (page 101) and the definition of "Linksquasikörper." Let us check the equivalence of (ii) and (iv): the law $(a + b) \circ c = a \circ c + b \circ c$ is automatic if $c \geqslant 0$; but for $c < 0$, $(a + b) \circ c = a \circ c + b \circ c$ if and only if $[\phi(a + b)] \cdot c = \phi(a) \cdot c + \phi(b) \cdot c$; the latter holds for all $c < 0$, $a, b \in F$ if and only if ϕ is additive.

COROLLARY 3. *A finite Moulton plane M must be (Y_∞, Y_∞)-transitive.*

Proof. Use Lemma 1 to represent M as $M_\phi(F)$, where $\phi(0) = 0$, $\phi(1) = 1$. By the Theorem of Carlitz, ϕ is an automorphism—in particular, it is additive.

COROLLARY 4. *If $M_\phi(F)$ denotes a finite Moulton plane with $\phi(0) = 0$, $\phi(1) = 1$, Conditions* (i)–(iv) *of Theorem* 5 *all hold in $M_\phi(F)$.*

Proof. The additivity of ϕ implies the remaining conditions.

Remark. Examples already given show that the conditions of Theorem 5 are not valid in every Moulton plane.

THEOREM 6. *If $\phi(0) = 0$ and $\phi(1) = 1$, a Moulton plane $M_\phi(F)$ determines a Cartesian group with associative multiplication if and only if Desargues' $(X_\infty, \{x = 0\})$-Theorem holds. Associativity of multiplication and the right-distributive law $c \circ (a + b) = c \circ a + c \circ b$ together are equivalent to Desargues' $(Y_\infty, \{y = 0\})$-Theorem.*

Proof. Since $M_\phi(F)$ satisfies $T(m, x, b) = m \circ x + b$ ("erste Zerlegbarkeitsbedingung"), the first part of *Satz 45* reduces to an equivalence between the associativity of multiplication and the $(X_\infty, \{x = 0\})$-Theorem. The second part of *Satz 45* becomes the final statement of Theorem 6, since the right-distributive law is equivalent to $T(m, x, m \circ b) = m \circ (x + b)$ ("zweite Zerlegbarkeitsbedingung") when $T(m, x, b) = m \circ x + b$ (Satz 35).

THEOREM 7. *Let $+$ and $\circ$ determine the Cartesian group for a Moulton plane $M_\phi(F)$, where $\phi(0) = 0$, $\phi(1) = 1$. Then*

(i) $c > 0$ *implies* $(a \circ b) \circ c = a \circ (b \circ c)$, *for all* $a, b, c \in F$.

(ii) $c < 0$ *and* $b > 0$ *imply* $(a \circ b) \circ c = a \circ (b \circ c)$ *for all* $a \in F$ *if and only if* ϕ *is multiplicative on* F.

(iii) $c < 0$ *and* $b < 0$ *imply* $(a \circ b) \circ c = a \circ (b \circ c)$ *for all* $a \in F$ *if and only if* $\phi(a) \cdot b = \phi^{-1}(a \cdot \phi(b))$.

Proof. Note first that ϕ preserves sign, since $\phi(0) = 0$ and ϕ is order-preserving.

(i) $(a \circ b) \circ c = (a \circ b) \cdot c = \lambda(a) \cdot bc = a \circ (bc) = a \circ (b \circ c)$, where $\lambda = \phi$ or $\mathcal{I}$ (the identity) according as $b < 0$ or $b > 0$.

(ii) $(a \circ b) \circ c = \phi(ab) \cdot c$, and $a \circ (b \circ c) = \phi(a) \cdot \phi(b) \cdot c$ [the multiplicative property being exactly what we need].

(iii) $(a \circ b) \circ c = \phi\{\phi(a) \cdot b\} \cdot c$, and $a \circ (b \circ c) = a \cdot \phi(b) \cdot c$, whence the condition $\phi(a) \cdot b = \phi^{-1}(a \cdot \phi(b))$.

COROLLARY 7. *Under the hypotheses of Theorem* 7, *the operation* $\circ$ *is associative if and only if* ϕ *is multiplicative and* $\phi = \phi^{-1}$.

COROLLARY 8. *Under the same hypotheses, the Desarguesian* $(X_\infty, \{x = 0\})$-*Theorem is valid in* $M_\phi(F)$ *if and only if* ϕ *denotes a multiplicative function of order* 2.

Proof. This follows from Theorems 6 and 7, and Corollary 7.

4. Collineations and isomorphisms on $M_\phi(F)$**.** A sequel to this paper will prove that some Moulton planes support a rather large group of collineations. It will treat isomorphisms between Moulton planes, and will show that a large class of "new" planes is obtained from the construction.

REFERENCES

1. R. Baer, *Homogeneity of projective planes*, Amer. J. of Math., *64* (1942), 137–152.
2. G. Birkhoff, *Lattice theory*, Colloq. Publ., Amer. Math. Soc., *25*, rev. ed. (1948).
3. L. Carlitz, *A theorem on permutations in a finite field*, Proc. Amer. Math. Soc., *2* (1960), 456–459.
4. R. F. Carmichael, *Introduction to the theory of groups of finite order*, Ginn and Co. (1937).
5. L. E. Dickson, *Linear groups with an exposition of the Galois theory*, B. G. Teubner (Leipzig, 1901).
6. M. Hall, Jr., *Projective planes*, Trans. Amer. Math. Soc., *54* (1943), 229–277.
7. ——— *The theory of groups*, Macmillan (New York, 1959).
8. D. Hilbert, *Grundlagen der Geometrie*, Teubner Verlag, Stuttgart 8. Auflage (1956).
9. P. Kustaanheimo, *On the relation of order in finite geometries*, Rend. Mat. e Appl. (5) *16* (1957), 292–296.
10. ——— *On the relation of congruence in finite geometries*, Rend. Mat. e Appl. (5) *16* (1957), 286–291.
11. C. Longo, *Teorema di Desargues ed omologie speciali in un piano grafico proiettive*, Lincei—Rend. Sc. fis. mat. e nat.—Vol. XXIV (Aprile, 1958).

12. F. R. Moulton, *A simple non-Desarguesian plane*, Trans. Amer. Math. Soc., *3* (1902), 192–195.
13. G. Pickert, *Projektive Ebenen*, Springer-Verlag (Berlin, 1955).
14. ——— *Nichtkommutative cartesische Gruppen*, Arch. Math., *3* (1952), 335–342.
15. K. Reidemeister, *Vorlesungen über Grundlagen der Geometrie* (Berlin, 1930).
16. E. Sperner, *Beziehungen zwischen geometrischer und algebraischer Anordnung*, Arch. Math., *1* (1948), 148–153.
17. O. Veblen and J. H. M. Wedderburn, *Non-Desarguesian and non-Pascalian geometries*, Trans. Amer. Math. Soc., *8* (1907), 379–388.

Syracuse University
Syracuse 10, New York

TETRASPHERES. I

A. DE MAJO

Les propriétés anallagmatiques de groupes de sphères ont été étudiées dans des contextes divers ces dernières années (voir **(4)**, **(5)**). Dans la note qui suit nous étudions les propriétés anallagmatiques de certains groupes de 4 sphères, nous plaçant à un point de vue élémentaire.

A fin de ne pas alourdir la rédaction nous omettrons de spécifier chaque fois que les points, droites ou sphères considérés sont toujours supposés être dans la position relative la plus générale possible compatible avec les définitions données.

L'index i pourra toujours prendre les valeurs $i = 1, 2, 3, 4$.

1. La configuration C_{12}**.** L'on sait (voir p.ex. **(3)** p. 134) que étant données 4 sphères en position générale les centres et axes de similitude de ces sphères forment une configuration de Reye.

Nous appelleront C_{12} une telle configuration, et rappelons qu'elle est projectivement équivalente à un cube, son centre et ses trois sommets à l'infini. On peut également considérer cette figure comme formée de trois tétraèdres tels que deux d'entre eux sont homologiques par rapport à un sommet et la face opposée du troisième. Chaque arête de l'un de ces tétraèdres coupe une arête de chacun des deux autres, et les deux points ainsi obtenus forment un quaterne harmonique avec les sommets du 1er tétraèdre sur cette arête. On obtient ainsi un groupe de 12 points, formant une C_{12}, que nous dirons *associée* à la 1ère. Notons que l'associativité est une propriété symétrique.

2. Octades. Nous appelleront octade la figure formée par quatre couples de points, A_i, A^i, sommets de quatres arêtes qui concourent au centre de l'octade.

Les 24 droites joignant les 8 points donnés 2 à 2 et ne contenant pas d'arête se coupent en 12 points, formant une C_{12}. Soit C_{12}' la configuration associée. Si l'on dénote par A_{ab} le point d'intersection des droites $A_aA_b - A^bA^a$, par A^{ab} celui des droites $A_aA^b - A_bA^a$, les trois tétraèdres associés à C_{12}' ont des couples d'arêtes opposées de la forme

$$\text{(I)}\quad A_{ab}A^{cd} - A_{cd}A^{ab},\quad \text{(II)}\quad A_{ab}A_{cd} - A^{ab}A^{cd},\quad \text{(III)}\quad A_{ab}A^{ab} - A_{cd}A^{cd}.$$

Les 32 droites joignant chacun des 4 sommets d'un tétraèdre du type I ou II aux 8 sommets de l'octade donnée concourent 4 par 4 aux 8 sommets d'une autre octade de même centre.

Reçu le 10 janvier, 1960.

On obtient en tout trois octades de cette manière et l'on prouve facilement la

Proposition 1. *A chaque octade correspondent deux autres octades de même centre, et ces trois octades sont associées à une C_{12}, de telle façon que chacun des sommets de l'une des octades est aligné avec chacun des sommets de chacune des deux autres octades par rapport aux sommets de l'un des tétraèdres associés à la C_{12}.*

Comme les sommets de deux telles octades se correspondent de 4 manières différentes de façon que les 8 droites joignant les sommets correspondants soient concourantes, l'on démontre facilement que pour que les figures formées par les sommets de deux octades associées soient inverses l'une de l'autre par rapport à 4 pôles différents il faut que:

(a) Les produits $OA_i \cdot OA^i$ soient égaux entre eux, ce qui implique que chacun des 4 couples de sommets opposés A_iA^i soit commun à 3 d'entre 4 sphères.

(b) Deux quelconques de ces 4 sphères se coupent suivant un angle égal ou supplémentaire à celui des deux autres.

Nous établirons au § 10 que ces deux conditions sont non seulement nécessaires mais aussi suffisantes.

3. Quadrisphères. Nous appellerons quadrisphère (S) la figure formée par 4 sphères S.

Trois quelconques de ces sphères ont en commun un couple de points. Le quadrisphère a donc 4 couples de sommets opposés A_iA^i, formant une octade. Les deux sommets opposés A_iA^i, alignés avec le centre radical O, sont inverses l'un de l'autre par rapport à la sphère S_0 orthogonale aux 4 sphères données.

L'inverse d'un quadrisphère est un autre quadrisphère, dont chaque angle est égal ou supplémentaire à l'angle correspondant de (S), et ceci d'après l'emplacement du pôle d'inversion.

Une étude détaillée des possibilités correspondantes montre que:

Proposition 2. *L'inversion ne peut donner pour les quadrisphères inverses que huit dispositions angulaires différentes.*

4. Quadrisphères de rayons donnés. Cherchons les points qui pris comme pôles d'inversion transforment le quadrisphère (S), dont les rayons sont r_i en un quadrisphère (S') dont les 4 rayons sont proportionnels à 4 nombres donnés, e_i. Chacun de ces pôles est commun à 6 surfaces, lieux des points tels que

$$\frac{w_a}{w_b} = \frac{r_a e_q}{r_b e_s}$$

(w_a = puissance par rapport à S_a). Or ce lieu se compose de deux sphères appartenant au faisceau linéaire défini par S_a et S_b. Nous dirons que la

sphère S_{ab}^{ei} dont le centre est extérieur aux points $O_a O_b$ (centres des sphères S_a et S_b) est extérieure.

Chacun des pôles cherchés peut être ainsi entièrement défini par la condition d'être commun à trois des 6 surfaces en question, par exemple:

$$S^e_{ab}\, pq\, S^i_{ab}\, pq - S^e_{ac}\, pr\, S^i_{ac}\, pr - S^e_{ad}\, ps\, S^i_{ad}\, ps.$$

On obtient ainsi 8 couples de points communs à 3 sphères orthogonales à S_0, donc inverses l'un de l'autre par rapport à S_0.

On montre facilement que les centres des sphères $S_{ij}{}^{in}$ forment une C_{12}.

L'étude du cas où les points cherchés sont tous réels est particulièrement intéressante, et l'on obtient alors la

PROPOSITION 3. *Dans le cas où les* 8 *couples de points pôles des inversions cherchées sont tous réels, ces couples correspondent biunivoquement aux* 8 *dispositions angulaires que l'on peut obtenir par inversion.*

Si l'on impose au quadrisphère inverse uniquement la condition d'avoir ses rayons proportionnels aux nombres e_i, sans considérer l'ordre des sphères correspondantes le nombre des pôles possibles est multiplié par 24, et l'on a:

PROPOSITION 4. *D'un quadrisphère l'on peut déduire* 384 *autres quadrisphères inverses du premier et tels que les rayons des sphères qui les composent soient proportionnels à* 4 *nombres donnés.*

5. Quadrisphères inverses égaux. Il est naturel de se demander combien, parmi ces 384 quadrisphères peuvent être égaux entre eux. L'on vérifie aisément que deux pôles inverses par rapport à la sphère S_0 donnent toujours par inversion des figures égales entre elles. Il faut donc examiner sous quelles conditions deux couples différents parmi les 192 couples de figures inverses pourront être composés de quadrisphères tous égaux. Ils devront avoir leurs sphères correspondantes égales entre elles, et aussi la même disposition angulaire.

Nous chercherons d'abord sous quelles conditions 2 couples de quadrisphères inverses ont la même disposition angulaire, ce qui implique soit l'égalité soit la symétrie des tétraèdres formés par leurs sommets. (On verra au § 11 que le deuxième cas ne se présente jamais).

Une étude détaillée des divers cas montre que si aucun des 6 angles n'est droit il faut que certains d'entre eux soient égaux ou supplémentaires aux autres. Le cas le plus intéressant est celui du

6. Tétrasphère. Nous appellerons ainsi le quadrisphère ayant trois angles arbitraires, les trois angles opposés étant égaux ou supplémentaires aux premiers.

Le tétrasphère sera dit pair si les couples d'angles opposés sont formés

d'angles égaux, impair s'ils se composent d'angles supplémentaires aux premiers.

Appliquant les résultats des sections précédentes on démontre aisément la

PROPOSITION 5. *Tout tétrasphère peut être transformé de 8 façons différentes en 8 tétrasphères de rayons donnés, égaux entre eux, dans des inversions par rapport à 8 sphères principales, ayant pour centres les sommets d'un même quadrisphère—il existe 8 tels quadrisphères, correspondant chacun à une disposition angulaire différente. Toutes les sphères considérées sont orthogonales à une même sphère.*

7. Orthosphère, équisphère, isosphère.

(a) Un tétrasphère dont tous les angles sont droits sera dit *orthosphère.* Il y a 192 couples de pôles d'inversion par rapport auxquels l'on peut transformer un orthosphère en 384 orthosphères égaux entre eux. Citons quelques propriétés de l'orthosphère:

PROPOSITION 6. *Les 4 sphères d'un orthosphère et sa sphère orthogonale forment un ensemble de 5 sphères 2 à 2 orthogonales et dont les centres sont les sommets d'un pentagone orthique* (*voir* **(1)**, **(2)**).

PROPOSITION 7. *L'orthosphère est invariant par rapport à une inversion dont la sphère principale est l'une de ces 5 sphères.*

(b) Si les 4 nombres e_i sont égaux entre eux, les figures inverses d'un quadrisphère obtenues comme au § 4 ont leur 4 sphères égales entre elles etseront appelées *équisphères.*

On voit facilement que dans ce cas le centre des 12 sphères du type $S_{ij}{}^{m}$ ne sont autres que les centres d'homotéthie des 4 sphères S_i, et ces 12 sphères sont donc les sphères bissectrices des couples de sphères du quadrisphère donné.

Dans ce cas ci il n'y a que 8 couples distincts de pôles d'inversion, correspondant chacun à une disposition angulaire différente.

(c) Les équisphères obtenus à partir d'un tétrasphère sont dénommés *isosphères.* Parmi leur nombreuses propriétés nous citons:

PROPOSITION 8. *Le tétraèdre ayant pour sommets les centres des sphères d'un iosphère est isofacial s'il est pair, orthocentrique s'il est impair.*

8. Orthosphère adjoint. Parmi les 8 couples de points définis au § 7(b) quatre sont des pôles d'inversions qui transforment le tétrasphère en un isosphère pair, nous les désignerons par D_iD^i et le quadrisphère dont ils sont les sommets par (D).

Si l'on soumet la figure formée par un tétrasphère, ses sphères bissectrices et le quadrisphère (D) à une inversion dont l'un des D_i est un pôle l'on

obtient comme transformé de (D) un orthosphère formé de 3 plans et une sphère. Il s'ensuit que (D) est un orthosphère.

On vérifie d'ailleurs aisément que les centres des sphères de (D) forment un tétraèdre conjugué à S_0.

On a donc la

PROPOSITION 9.

(a) *Les* 12 *sphères bissectrices d'un tétrasphère se coupent* 6 *à* 6 *en* 16 *points parmi lesquels* 8 *sont les sommets d'un orthosphère, dit adjoint au tétrasphère.*

(b) *Les centres des sphères de cet orthosphère forment un tétraèdre orthocentrique conjugué à la sphère orthogonale du tétrasphère.*

9. Propriété fondamentale du tétrasphère. Il s'ensuit des résultats du § 6 que tout point de l'espace peut être pris comme sommet d'un quadrisphère α tel que tous ses sommets donnent par inversion d'un tétrasphère donné des figures égales.

Comme l'orthosphère adjoint à un tétrasphère s'en déduit par des opérations anallagmatiques, les figures inverses de cet orthosphère par rapport à ces mêmes points seront également égales, et étant donné l'un de ces points la construction des autres sommets peut se faire à partir de (D) et non de (S); de sorte que les quadrisphères comme α ne dépendent que de (D). Un tel quadrisphère sera dit *annexé* à (D).

D'autre part l'on vérifie aisément que tout tétrasphère joint à (D) (c.à.d. tel que (D) soit son adjoint) est également annexé à (D). Or tout point de l'espace est l'un des centres des sphères d'un tétraèdre joint à un orthosphère donné. Deux tétrasphères joints à un même orthosphère sont dits *associés.*

Nous conclurons:

PROPOSITION 10.

(a) *Etant donné un tétrasphère, tout point de l'espace est sommet d'un second tétrasphère, associé au premier, tel que tous les sommets de l'un quelconque des deux sont des pôles d'inversions transformant l'autre en tétrasphères égaux.*

(b) *La propriété d'être associé est transitive pour les tétrasphères.*

10. Tétrasphères conjugués. Au § 8 nous avons considéré l'orthosphère (D) dont les sommets sont 4 couples de pôles transformant le tétrasphère en isosphère pair, les 8 pôles transformant le tétrasphère en isosphère impair sont les sommets d'un autre quadrisphère, dénommé (E).

Nous appelerons P_i les centres des sphères de l'orthosphère adjoint à un tétrasphère.

Considérons la figure formée par un tétrasphère (S), ses sphères bissectrices, les quadrisphères (D) et (E), et soumettons-la à une inversion de pôle P_i. Il est facile de voir que la figure est invariante, pour une puissance d'inversion convenable.

Puisque (E) est invariant, que les sphères bissectrices sont transformées en sphères bissectrices, et que (S) n'est pas invariant en général (sauf si ses 4 centres sont coplanaires), il s'ensuit que (S) et (E) sont inverses l'un de l'autre; donc (E) est également un tétrasphère. D'où la

PROPOSITION 11. *(S) et (E) sont deux tétrasphères inverses l'un de l'autre par rapport à chacun des 4 pôles P_i. Nous dirons qu'ils sont conjugués.*

Ceci démontre que les conditions nécessaires énoncées au § 2 pour que deux octades soient inverses l'une de l'autre sont aussi suffisantes.

11. Egalité ou symétrie. Les 4 poles P^i, pieds des hauteurs du tétraèdre P formé par les points P_i transforment également (S) en des tétrasphères égaux à (E), mais ne coïncident pas avec ce dernier, car ils lui sont symétriques par rapport aux hauteurs P_iP^i.

Ceci nous permet de montrer que quels que soient les 8 pôles déduits par continuité des pôles P_i et P^i les figures obtenues sont toujours égales et jamais symétriques—comme annoncé au § 5.

En effet l'égalité entre figures inverses ne pourrait venir à disparaître que si l'une des figures vient à posséder un plan de symétrie, ce qui exige que le pôle correspondant soit sur S_0. Or on vérifie aisément que deux pôles inverses par rapport à S_0 donnent des figures inverses égales, le passage d'un pôle par S_0 ne peut donc pas changer l'égalité en symétrie.

12. Construction de tétrasphères associés. La considération de l'isosphère associé à P dont les sommets sont les points P_i et P^i, et de son inverse (Q) par rapport à un pôle P_i, qui a pour sommets les 6 pieds Q_{ab} des perpendiculaires communes aux arêtes opposées de P, le point O et un point à l'infini permet de voir facilement que l'on a la

PROPOSITION 12. *Les inverses du tétrasphère (S) par rapport aux pieds des perpendiculaires communes aux arêtes de son tétraèdre associé P sont les symétriques de ce tétrasphère par rapport aux plans hauteurs menés par les arêtes opposées à ces pôles.*

Ceci nous permet de construire un tétrasphère associé à un tétraèdre orthocentrique, lorsque l'on se donne l'un de ses sommets, A_a p.ex.

A^a est l'inverse de A_a par rapport à S_0.

A_b, A^b sont à l'intersection des droites joignant Q_{ab} ou Q_{cd} aux symétriques de A_a et A^a par rapport aux plans hauteurs P_{cd} (mené par P_c et P_d) et P_{ab}.

La considération des questions de réalité nous mène alors à la

PROPOSITION 14. *Un tétrasphère réel, associé à un tétraèdre réel conjugué à une sphère S_0 imaginaire, a ses 8 sommets réels; lorsque la sphère S_0 est réelle, les 8 sommets de ce tétrasphère réel sont simultanément réels ou imaginaires.*

13. Tétrasphère et pentagone orthique. Au § 10 nous avons considéré le tétrasphère conjugué à (S). Soit B_i, B^i ses sommets.

L'ensemble de 16 points A_i, A^i, B_i, B^i jouit de nombreuses propriétés par rapport aux inversions dont les 5 sphères principales, 2 à 2 orthogonales, sont la sphère S_0 et les 4 sphères principales D_i, de centres P_i.

Le tableau ci-après donne pour chacune de ces sphères la répartition des 16 sommets entre les 10 tétrasphères deux à deux conjugués.

TABLEAU

Sphère orthogonale	Centre radical	Premier tétrasphère	Second tétrasphère
S_0	O	$A_1A^1-A_2A^2-A_3A^3-A_4A^4$	$B_1B^1-B_2B^2-B_3B^3-B_4B^4$
D_1	P_1	$A_1B_1-A_2B^2-A^3B^3-A^4B^4$	$A^1B^1-A^2B_2-A_3B_3-A_4B_4$
D_2	P_2	$A_1B_2-A_2B^1-A_3B^4-A_4B^3$	$A^1B^2-A^2B_1-A^3B_4-A^4B_3$
D_3	P_3	$A_1B_3-A_2B^4-A^3B^1-A_4B^2$	$A^1B^3-A^2B_1-A_3B_1-A^4B_2$
D_4	P_4	$A_1B_4-A_2B^3-A_3B^2-A^4B^1$	$A^1B^4-A^2B_3-A^3B_2-A_4B_1$

BIBLIOGRAPHIE

1. N. A. Court, *On five mutually orthogonal spheres*, Ann. of Math., *30* (1929).
2. ——— *Sur quatre sphères réelles deux à deux orthogonales*, Mathesis, *65* (1956).
3. Hilbert, Cohn, Vossen, *Geometry and the imagination*, (New York, 1952).
4. R. Lagrange, *Produits d'inversions et métrique conforme* (Paris, 1957).
5. ——— *Sur les systèmes isogonaux de shpères*, Ann. Scient. E. N. Sup., t.73 (1959).

Paris

POLAR MEANS OF CONVEX BODIES AND A DUAL TO THE BRUNN-MINKOWSKI THEOREM

WILLIAM J. FIREY

1. Introduction. This paper deals with processes of combining convex bodies in Euclidean n-space which are, in a sense, dual to the process of Minkowski addition and some of its generalizations.

All the convex bodies considered will have a common interior point Q. Variables x and y denote vectors drawn from Q; we shall speak of their terminal points as the points x and y. Unit vectors will be denoted by u; $||x||$ signifies the length of x. Convex bodies will be symbolized by K with distinguishing marks. ∂K means the boundary of K. λK will mean the image of K under a homothetic transformation in the ratio $\lambda : 1$. The centre of the homothety will always be Q.

The distance function $F(x)$ of a convex body is defined as follows: let y be the vector having the same direction as x which terminates at ∂K, then $F(x) = ||x||/||y||$. If $x = 0$, we set $F(0) = 0$. The points x of K satisfy $F(x) \leqq 1$ with equality if and only if x is a point of ∂K. Let $u = x/||x||$; then $\rho = 1/F(u) = f(u)$ is the polar co-ordinate equation of ∂K with respect to a co-ordinate system with pole at Q. Since Q is an interior point of K, $F(u)$ is continuous and bounded.

The distance function satisfies: (a) $F(x) > 0$ for $x \neq 0$, $F(0) = 0$; (b) $F(\mu x) = \mu F(x)$ for $\mu > 0$; (c) $F(x + y) \leqq F(x) + F(y)$ for any two vectors x and y. Conversely, any function $F(x)$ satisfying (a) through (c) is the distance function of a unique convex body K (cf. **(1)**, p. 22).

The following observations regarding distance functions should be borne in mind; they follow immediately from the definition. $F_0(x) \geqq F_1(x)$ if and only if $K_0 \subseteq K_1$. If the distance function of K is $F(x)$, that of λK is $F(x)/\lambda$.

If $F_i(x)$, $(i = 0, 1)$, is the distance function of the body K_i containing Q as an interior point, then

$$F_\vartheta^{(1)}(x) = (1 - \vartheta)F_0(x) + \vartheta F_1(x), \qquad 0 \leqq \vartheta \leqq 1,$$

and, more generally,

$$F_\vartheta^{(p)}(x) = \sqrt[p]{[(1 - \vartheta)F_0^p(x) + \vartheta F_1^p(x)]}, \qquad 1 \leqq p \leqq \infty,$$

satisfy conditions (a) through (c). By $F_\vartheta^{(\infty)}(x)$ we mean

$$\lim_{p \to \infty} F_\vartheta^{(p)}(x) = \max(F_0(x), F_1(x))$$

Received March 28, 1960.

for $0 < \vartheta < 1$ with $F_i^{(\infty)}(x) = F_i(x)$. Conditions (a) and (b) are obviously satisfied. Condition (c) is a consequence of Minkowski's inequality. Let $a_i = b_i + c_i$; Minkowski's inequality is

$$\sqrt[p]{[(1-\vartheta)a_0^p + \vartheta a_1^p]} \leqq \sqrt[p]{[(1-\vartheta)b_0^p + \vartheta b_1^p]} + \sqrt[p]{[(1-\vartheta)c_0^p + \vartheta c_1^p]}.$$

If $a_i \leqq b_i + c_i$, the inequality is clearly still valid. Set $a_i = F_i(x+y)$, $b_i = F_i(x)$ and $c_i = F_i(y)$ and condition (c) is verified for $F_\vartheta^{(p)}$. A limit argument establishes (c) for $p = \infty$. Consequently we may speak of a unique convex body $\dot{K}_\vartheta^{(p)}$ having the distance function $F_\vartheta^{(p)}$. We will call this body the pth dot-mean of K_0 and K_1. It clearly contains Q as an interior point. For $1 \leqq p < \infty$, the body

$$\sqrt[p]{2}\dot{K}_{1/2}^{(p)}$$

will be denoted by $\dot{S}^{(p)}(K_0, K_1)$ and called the pth dot-sum of K_0 and K_1. Its distance function is $\sqrt[p]{[F_0^p(x) + F_1^p(x)]}$. We set

$$\dot{S}^{(\infty)}(K_0, K_1) = \dot{K}_{1/2}^{(\infty)}.$$

We obtain a direct geometric meaning for $\dot{K}_\vartheta^{(p)}$ as follows. If the polar co-ordinate equation of ∂K_i is $\rho = f_i(u)$, then the polar co-ordinate equation of $\partial\dot{K}_\vartheta^{(p)}$ is

$$\rho = 1 \Big/ \sqrt[p]{\left[\frac{(1-\vartheta)}{f_0^p(u)} + \frac{\vartheta}{f_1^p(u)}\right]} \qquad \text{for } 1 \leqq p < \infty,$$

$$\rho = \min\,(f_0(u), f_1(u)) \qquad \text{for } p = \infty.$$

In particular if $p = 1$, ρ is the harmonic mean of the distances to ∂K_0 and ∂K_1 in the direction u.

$$\dot{K}_\vartheta^{(\infty)} = K_0 \cap K_1$$

for $0 < \vartheta < 1$.

In § 2, we first take up some elementary rules about such combinations of convex bodies. A deviation or metric in a space of convex bodies is introduced. The duality mentioned at the beginning of the paper is discussed and with its aid, we examine the topology induced by the deviation measure.

Section 3 is devoted to the dependence of the family $\{\dot{K}_\vartheta^{(p)}\}$ on K_0, K_1 and the parameters p and ϑ, for $1 \leqq p < \infty$. The dependence is continuous; the family is monotonic decreasing in p and concave with respect to ϑ. The special case $p = \infty$ is considered separately.

We establish a theorem of the Brunn–Minkowski type for the family $\{\dot{K}_\vartheta^{(p)}\}$ in the final section. This is

$$V^{1/n}(\dot{K}_\vartheta^{(p)}) \leqq 1/\sqrt[p]{[(1-\vartheta)V^{-p/n}(K_0) + \vartheta V^{-p/n}(K_1)]} \qquad \text{for } 1 \leqq p < \infty,$$

$$V(\dot{K}_\vartheta^{(\infty)}) \leqq \min(V(K_0), V(K_1)) \qquad \text{for } 0 < \vartheta < 1.$$

Here $V(K)$ signifies the volume of the convex body K.

A discussion of the cases of equality is included.

2. Measures of deviation. The following rules follow immediately from the properties of $S_p(a_0, a_1) = \sqrt[p]{[a_0^p + a_1^p]}$ for non-negative numbers a_i applied to the appropriate distance functions.

(i) $\dot{S}^{(p)}(\lambda K_0, \lambda K_1) = \lambda \dot{S}^{(p)}(K_0, K_1)$.

(ii) $\dot{S}^{(p)}(K_0, K_1) = \dot{S}^{(p)}(K_1, K_0)$.

(iii) $\dot{S}^{(p)}(\dot{S}^{(p)}(K_0, K_1), K_2) = \dot{S}^{(p)}(K_0, \dot{S}^{(p)}(K_1, K_2))$.

This last rule allows us to write without misunderstanding $\dot{S}^{(p)}(K_0, K_1, \ldots, K_m)$ defined inductively as

$$\dot{S}^{(p)}(\dot{S}^{(p)}(K_0, K_1, \ldots, K_{m-1}), K_m).$$

In turn we set

$$\dot{S}^{(p)}(\sqrt[p]{w_0}K_0, \sqrt[p]{w_1}K_1, \ldots, \sqrt[p]{w_m}K_m) = \dot{M}^{(p)}(K_0, K_1, \ldots, K_m)$$

if

$$\sum_{i=1}^{m} w_i = 1, w_i \geqq 0, 1 \leqq p < \infty.$$

$\dot{M}^{(p)}(K_0, K_1) = \dot{K}_\vartheta^{(p)}$ with $\vartheta = w_1$. We define $\dot{M}^{(\infty)}(K_0, K_1, \ldots, K_m)$ and $\dot{S}^{(\infty)}(K_0, K_1, \ldots, K_m)$ as bodies whose distance functions are

$$\lim_{p\to\infty} M_p(F_0, F_1, \ldots, F_m), \lim_{p\to\infty} S_p(F_0, F_1, \ldots, F_m).$$

Since these limits are equal $\dot{M}^{(\infty)}(K_0, K_1, \ldots, K_m)$, $\dot{S}^{(\infty)}(K_0, K_1, \ldots, K_m)$ are the same body. This is the convex body whose distance function is max $(F_0, F_1, \ldots, F_m)$. $\partial \dot{M}^{(\infty)}(K_0, K_1, \ldots, K_m)$ has the polar co-ordinate equation $\rho = \min(f_0, f_1, \ldots, f_m)$ if ∂K_i has the equation $\rho = f_i(u)$. Clearly

$$\dot{M}^{(\infty)}(K_0, K_1, \ldots, K_m) = K_0 \cap K_1 \cap \ldots \cap K_m.$$

We always have $\dot{S}^{(p)}(K_0, K_1) \subset K_i$ since

$$\sqrt[p]{[F_0^p(x) + F_1^p(x)]} > F_i(x)$$

for $x \neq 0$.

The bodies $\dot{S}^{(p)}(K_0, K_1)$ and $\dot{K}_\vartheta^{(p)}$ are not translation-invariant in the sense displayed by the usual Minkowski sum $K_0 + K_1$. In the case of Minkowski sums, if K_i is translated by the addition of a vector t_i to each vector in K_i, then $K_0 + K_1$ is translated by the addition of the vector $t_0 + t_1$. It can be proved that, in general, there is no such translation vector for $\dot{S}^{(p)}(K_0, K_1)$ or $\dot{K}_\vartheta^{(p)}$. For this reason we must distinguish bodies which differ by a translation.

A measure of deviation between the two convex bodies is defined as follows. Let E be the sphere of radius one, centred at Q. For $1 \leqq p < \infty$, consider those numbers $\lambda > 0$ such that $\dot{S}^{(p)}(K_0, \lambda E) \subseteq K_1$ and $\dot{S}^{(p)}(K_1, \lambda E) \subseteq K_0$. We define $\dot{\delta}^{(p)}(K_0, K_1)$ to be the greatest lower bound of the numbers $1/\lambda$. In terms

of distance functions, if $F_i(x)$ is the distance function of K_i, $\dot{\delta}^{(p)}(K_0, K_1)$ is the greatest lower bound of numbers $1/\lambda = \mu$ such that

$$\sqrt[p]{[F_0^p(x) + \mu^p||x||^p]} \geqq F_1(x)$$

and

$$\sqrt[p]{[F_1^p(x) + \mu^p||x||^p]} \geqq F_0(x).$$

Since such function $F_i(x)$ is continuous and bounded over $||x|| = 1$, we have

$$\dot{\delta}^{(p)}(K_0, K_1) = \max \sqrt[p]{|F_0^p(u) - F_1^p(u)|},$$

the maximum being taken over the sphere of directions u. Clearly $\dot{\delta}^{(p)}(K_0, K_1) \geqq 0$ with equality if and only if $F_0(x) \equiv F_1(x)$, that is $K_0 = K_1$. Further $\dot{\delta}^{(p)}(K_0, K_1) = \dot{\delta}^{(p)}(K_1, K_0)$. The deviation satisfies a triangle inequality:

$$\dot{\delta}^{(p)}(K_0, K_2) \leqq \dot{\delta}^{(p)}(K_0, K_1) + \dot{\delta}^{(p)}(K_1, K_2).$$

For let

$$\mu_1 = \dot{\delta}^{(p)}(K_0, K_1),$$
$$\mu_2 = \dot{\delta}^{(p)}(K_0, K_2),$$
$$\mu_3 = \dot{\delta}^{(p)}(K_1, K_2).$$

Then

$$\mu_2 = \max \sqrt[p]{|F_0^p(u) - F_2^p(u)|} \leqq \max \sqrt[p]{[|F_0^p(u) - F_1^p(u)| + |F_1^p(u) - F_2^p(u)|]}$$
$$\leqq \max \sqrt[p]{|F_0^p(u) - F_1^p(u)|} + \max \sqrt[p]{|F_1^p(u) - F_2^p(u)|} = \mu_1 + \mu_3,$$

all the maxima being taken over the unit sphere of directions u.

For $p = \infty$, we define $\dot{\delta}^{(\infty)}(K_0, K_1)$ to be

$$\max_{||u||=1} \left(\max_{(0,1)} [F_0(u), F_1(u)]\right)$$

if K_0 and K_1 are not identical and take $\dot{\delta}^{(\infty)}(K_0, K_0) = 0$. $\dot{\delta}^{(\infty)}(K_0, K_1)$ is thus the reciprocal of the radius of the largest sphere centred at Q which lies in $K_0 \cap K_1$. We may alternately describe $\dot{\delta}^{(\infty)}(K_0, K_1)$ as $\max(1/\nu_0, 1/\nu_1)$ where $\nu_i E$ is the largest sphere centred at Q contained in K_i. Clearly $\dot{\delta}^{(\infty)}(K_0, K_1) = \dot{\delta}^{(\infty)}(K_1, K_0)$ and $\dot{\delta}^{(\infty)}(K_0, K_1) \geqq 0$ with equality if and only if $K_0 = K_1$. This deviation satisfies a triangle inequality:

$$\dot{\delta}^{(\infty)}(K_0, K_2) \leqq \dot{\delta}^{(\infty)}(K_0, K_1) + \dot{\delta}^{(\infty)}(K_1, K_2).$$

If $K_0 = K_2$, this follows from the non-negativity of the deviation. If $K_0 = K_1$ or $K_1 = K_2$, there is obvious equality. Otherwise, using the numbers ν_0, ν_1, ν_2 defined above, we have

$$\max\left(\frac{1}{\nu_0}, \frac{1}{\nu_2}\right) \leqq \max\left(\frac{1}{\nu_0}, \frac{1}{\nu_1}, \frac{1}{\nu_2}\right) < \max\left(\frac{1}{\nu_0}, \frac{1}{\nu_1}\right) + \max\left(\frac{1}{\nu_1}, \frac{1}{\nu_2}\right)$$

which proves the assertion.

Thus, for $1 \leqq p \leqq \infty$, the deviations $\dot{\delta}^{(p)}(K_0, K_1)$ satisfy the requirements

for a metric in the space of convex bodies. For the remainder of the section, deviations will be considered only for $1 \leq p < \infty$.

Let K be a convex body with distance function $F(x)$. We denote by $\hat{K}$ the polar reciprocal of K with respect to the unit sphere E centred at Q. The support function with respect to Q of $\hat{K}$ is defined as follows. Let x be any point other than Q, z a vector from Q in the direction of x which terminates at the support plane of $\hat{K}$ normal to x. The support function of $\hat{K}$ is $||z||\cdot||x||$. Since K and $\hat{K}$ are polar reciprocals with respect to E, if y is the vector from Q having the same direction as x and terminating at ∂K, we have $||y||\cdot||z|| = 1$. Hence the support function of $\hat{K}$ is $||x||/||y|| = F(x)$. Further, if $H(x)$ is the distance function of $\hat{K}$, then $H(x)$ is the support function of K. If Q is an interior point of K, it is an interior point of $\hat{K}$. Consider the convex body $\dot{K}_\vartheta^{(p)}$; its polar reciprocal $\hat{\dot{K}}_\vartheta^{(p)}$ has

$$\sqrt[p]{[(1-\vartheta)F_0^p(x) + \vartheta F_1^p(x)]}$$

as its support function. This support function is the pth mean of the support functions of $\hat{K}_0$ and $\hat{K}_1$. In particular for $p = 1$, $\hat{\dot{K}}_\vartheta^{(p)}$ is the usual Minkowski mean $(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1$. More generally $\hat{\dot{K}}_\vartheta^{(p)}$ is the convex body denoted by $\hat{K}_\vartheta^{(p)}$ called the pth mean of $\hat{K}_0$, $\hat{K}_1$ in **(2)**. Similarly $\hat{\dot{S}}^{(p)}(K_0, K_1) = S^{(p)}(\hat{K}_0, \hat{K}_1)$.

It is convenient to express these notions in terms of the space $\mathcal{K}_p$ of convex bodies K with metric $\dot{\delta}^{(p)}$ and the space $\hat{\mathcal{K}}_p$ of convex bodies $\hat{K}$ with metric $\delta^{(p)}$ introduced in **(2)**. There $\delta^{(p)}(\hat{K}_0, \hat{K}_1)$ was defined as the greatest lower bound of numbers μ such that

$$\sqrt[p]{[F_0^p(x) + \mu^p||x||^p]} \geq F_1(x)$$

and

$$\sqrt[p]{[F_1^p(x) + \mu^p||x||^p]} \geq F_0(x)$$

where $F_i(x)$ is the support function of $\hat{K}_i$. Polar reciprocation with respect to E is an involutary mapping $R_p\colon \mathcal{K}_p \to \hat{\mathcal{K}}_p$. Under this mapping pth dot-means correspond to pth means.

We have directly from the definitions of $\dot{\delta}^{(p)}$ and $\delta^{(p)}$ that $\dot{\delta}^{(p)}(K_0, K_1) = \delta^{(p)}(\hat{K}_0, \hat{K}_1)$. Therefore R_p is a homeomorphism. In **(2)** it was shown that the metrics $\delta^{(p)}$ are topologically equivalent and so it follows also for the metrics $\dot{\delta}^{(p)}$.

We summarize.

THEOREM 1. *Polar reciprocation with respect to E furnishes a homeomorphism* $\mathcal{K}_p \to \hat{\mathcal{K}}_p$, *for* $1 \leq p < \infty$ *and for each such* p *and* q *satisfying* $1 \leq q < \infty$, $\mathcal{K}_p$ *is homeomorphic to* $\mathcal{K}_q$.

Let E_m $(1 \leq m < n)$ be an m-dimensional linear subspace of the Euclidean n-space which contains Q. The distance function of $K \cap E_m$ in E_m is the restriction of the distance function of K to vectors in E_m. Hence in E_m we have

$$\dot{S}^{(p)}(K_0, K_1) \cap E_m = \dot{S}^{(p)}(K_0 \cap E_m, K_1 \cap E_m).$$

This is the dual of the following result. Let K^* be the projection of K onto E_m; then

$$S^{(p)}(K_0^*, K_1^*) = [S^{(p)}(K_0, K_1)]^*.$$

We have further

$$S^{(p)}(K_0 \cap E_m, K_1 \cap E_m) \subseteq S^{(p)}(K_0, K_1) \cap E_m$$

and, as the dual of this result

$$\dot{S}^{(p)}(K_0^*, K_1^*) \supseteq [\dot{S}^{(p)}(K_0, K_1)]^*.$$

The latter follows from the former with the observations that if F^* is the support function of $\hat{K} \cap E_m$ then it is the distance function of K^*, and by the first inclusion

$$\sqrt[p]{[(F_0^*)^p + (F_1^*)^p]} \leq (\sqrt[p]{[F_0^p + F_1^p]})^*.$$

3. Dependence of the means on their parameters. The pth dot-means $\dot{K}_\vartheta^{(p)}$ depend continuously on p, ϑ, K_0 and K_1 in the following sense. Let S be the space of elements (p, ϑ, K_0, K_1) where $1 \leq p \leq P < \infty$, $0 \leq \vartheta \leq 1$, K_i in $\mathcal{K}_1$ with the distance $d(e, e')$ between elements $e = (p, \vartheta, K_0, K_1)$ and $e' = (p', \vartheta', K_0', K_1')$ defined as $|p - p'| + |\vartheta - \vartheta'| + \dot{\delta}^{(1)}(K_0, K_0') + \dot{\delta}^{(1)}(K_1, K_1')$. By Theorem 1, the deviation $\dot{\delta}^{(1)}$ can be replaced by any of the deviations $\dot{\delta}^{(q)}, \delta^{(q)}$ for finite $q \geq 1$. Further let $K(e)$ be the pth dot-mean $\dot{K}_\vartheta^{(p)}$ associated with element e. $K(e)$ is continuous in e, that is if $\{e_n\}$ is any sequence of elements of S for which

$$\lim_{n\to\infty} d(e_n, e) = 0,$$

we have

$$\lim_{n\to\infty} \dot{\delta}^{(1)}(K(e_n), K(e)) = 0.$$

To demonstrate this continuity, we first remark that the algebraic function

$$f(p, \vartheta, a_0, a_1) = \sqrt[p]{[(1-\vartheta)a_0^p + \vartheta a_1^p]}$$

has no singularities for (p, ϑ, a_0, a_1) satisfying $0 < A \leq a_i \leq B < \infty$, $0 \leq \vartheta \leq 1$, $1 \leq p \leq P < \infty$ and so is uniformly continuous for such (p, ϑ, a_0, a_1). Suppose that $\{F_{0n}(x)\}$ and $\{F_{1n}(x)\}$ converge to $F_0(x)$ and $F_1(x)$ uniformly for $||x|| = 1$ and further satisfy $A \leq F_{in}(x) \leq B$. Then it is easily shown that $\{f(p_n, \vartheta_n, F_{0n}(x), F_{1n}(x))\}$ is a sequence converging to $f(p, \vartheta, F_0(x), F_1(x))$ uniformly for $||x|| = 1$, where $\{p_n\}$ and $\{\vartheta_n\}$ converge to p and ϑ and satisfy $1 \leq p_n \leq P$, $0 \leq \vartheta_n \leq 1$.

The convergence of a sequence of elements $e_n = (p_n, \vartheta_n, K_{0n}, K_{1n})$ of S to element e of S implies

$$\lim_{n\to\infty} \dot{\delta}^{(1)}(K_i, K_{in}) = 0$$

which in turn is equivalent to the convergence of the associated sequences of distance functions $\{F_{in}(x)\}$ to $F_i(x)$ uniformly for $||x|| = 1$. Moreover, since all the bodies in the sequences $\{K_{in}\}$ as well as the limit bodies K_i are in $\mathscr{K}_1$ we know that there is a sphere $(1/A)E$ containing each K_i and K_{in}, and a sphere $(1/B)E$ contained in each K_i and K_{in}. From this it follows that $0 < A \leqq F_{in}(x) \leqq B < \infty$. Thus, by the preceding paragraph, the convergence of $\{e_n\}$ to e entails the convergence of $\{f(p_n, \vartheta_n, F_{0n}(x), F_{1n}(x))\}$ to $f(p, \vartheta, F_0(x), F_1(x))$ uniformly for $||x|| = 1$. This is to say that

$$\lim_{n\to\infty} \dot{\delta}^{(1)}(K(e_n), K(e)) = 0$$

as asserted.

We next examine inclusion relations among the means $\dot{K}_\vartheta{}^{(p)}$. Since

$$\sqrt[p]{[(1-\vartheta)F_0^p(x) + \vartheta F_1^p(x)]} \leqq \sqrt[q]{[(1-\vartheta)F_0^q(x) + \vartheta F_1^q(x)]}$$

for $1 \leqq p < q \leqq \infty$ with equality if and only if $F_0(x) = F_1(x)$, we have $\dot{K}_\vartheta{}^{(p)} \supseteqq \dot{K}_\vartheta{}^{(q)}$ with equality if and only if $K_0 = K_1$. Thus the means are either constant if $K_0 = K_1 = \dot{K}_\vartheta{}^{(p)}$ or are strictly monotonic decreasing in p from $\dot{K}_\vartheta{}^{(1)}$ to $K_0 \cap K_1$.

Finally consider the family $\{\dot{K}_\vartheta{}^{(p)}\}$ for fixed p and varying ϑ. For $p = \infty$, it is geometrically obvious that the family is convex by which we mean that

$$\dot{K}_{\vartheta'}^{(p)} \subseteq (1-\vartheta)\dot{K}_{\vartheta_0}^{(p)} + \vartheta\dot{K}_{\vartheta_1}^{(p)}$$

where $\vartheta' = (1-\vartheta)\vartheta_0 + \vartheta\vartheta_1$. But this is true for all p satisfying $1 \leqq p \leqq \infty$. In virtue of the monotonicity in p discussed in the preceding paragraph, it is enough to show the asserted convexity for $p = 1$.

We make a further reduction of the problem. Since

$$\hat{\dot{K}}_\vartheta^{(1)} = (1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1,$$

we have

$$\begin{aligned}\hat{\dot{K}}_{\vartheta'}^{(1)} &= [(1-\vartheta')\hat{K}_0 + \vartheta'\hat{K}_1]\hat{}\\ &= [(1-\vartheta)[(1-\vartheta_0)\hat{K}_0 + \vartheta\hat{K}_1] + \vartheta[(1-\vartheta_1)\hat{K}_0 + \vartheta_1\hat{K}_1]]\hat{},\end{aligned}$$

and

$$(1-\vartheta)\dot{K}_{\vartheta_0}^{(1)} + \vartheta\dot{K}_{\vartheta_1}^{(1)} = (1-\vartheta)[(1-\vartheta_0)\hat{K}_0 + \vartheta_0\hat{K}_1]\hat{} + \vartheta[(1-\vartheta_1)\hat{K}_0 + \vartheta_1\hat{K}_1]\hat{}.$$

Set $K = (1-\vartheta_0)\hat{K}_0 + \vartheta_0\hat{K}_1$ and $K' = (1-\vartheta_1)\hat{K}_0 + \vartheta_1\hat{K}_1$. In terms of K, K' we must prove that $[(1-\vartheta)K + \vartheta K']\hat{} \subseteq (1-\vartheta)\hat{K} + \vartheta\hat{K}'$.

On a ray r from Q let x be on ∂K, x' on $\partial K'$. Then $x_\vartheta = (1-\vartheta)x + \vartheta x'$ is a point, in general interior, of the Minkowski sum $(1-\vartheta)K + \vartheta K'$. Let Π, Π', Π_ϑ be the polar planes of x, x', and x_ϑ. These planes are orthogonal to r and meet r in points z, z', and z_ϑ. Π and Π' are support planes of K and K'. Π_ϑ is a plane exterior to $[(1-\vartheta)K + \vartheta K']\hat{}$ unless x_ϑ happens to be a boundary point of $(1-\vartheta)K + \vartheta K'$, in which case Π_ϑ is a support plane of $[(1-\vartheta)K$

$+ \vartheta K']^\wedge$. Let $\bar{z} = (1-\vartheta)z + \vartheta z'$. The plane Π, orthogonal to r through $\bar{z}$ is a support plane of $(1-\vartheta)\hat{K} + \vartheta\hat{K}'$.

If we can show that $z_\vartheta \leqq \bar{z}$, it will follow that $\bar{\Pi}$ is either exterior to $[(1-\vartheta)K + \vartheta K']^\wedge$ or coincides with Π_ϑ if $z_\vartheta = \bar{z}$. Since r is arbitrary, this will prove that

$$[(1-\vartheta)K + \vartheta K']^\wedge \subseteq (1-\vartheta)\hat{K} + \vartheta\hat{K}'.$$

We have from the polarity relations:

$$||z||\cdot||x|| = ||z'||\cdot||x'|| = ||z_\vartheta||\cdot||x_\vartheta|| = 1.$$

Hence

$$||z_\vartheta|| = \frac{\dfrac{(1-\vartheta)}{||z||}\cdot||z||\cdot||x|| + \dfrac{\vartheta}{||z'||}\cdot||z'||\cdot||x'||}{\dfrac{(1-\vartheta)}{||z||}\cdot||x_\vartheta|| + \dfrac{\vartheta}{||z'||}\cdot||x_\vartheta||}$$

$$= \frac{||(1-\vartheta)x + \vartheta x'||}{||x_\vartheta||\cdot\left(\dfrac{1-\vartheta}{||z||} + \dfrac{\vartheta}{||z'||}\right)}.$$

In the last step, we have utilized the collinearity of Q, x, and x'. Continuing:

$$||z_\vartheta|| = \frac{1}{\dfrac{(1-\vartheta)}{||z||} + \dfrac{\vartheta}{||z'||}} \leqq (1-\vartheta)||z|| + \vartheta||z'|| = ||\bar{z}||$$

where the collinearity of Q, z, z', and z_ϑ has been used. In the inequality of the arithmetic and harmonic means, there is equality if and only if $||z|| = ||z'||$, from which we conclude that the original inclusion is an equality if and only if $K = K'$.

This argument proves the convexity of $\{\dot{K}_\vartheta^{(p)}\}$. The family is linear if and only if

$$\dot{K}_{\vartheta_0}^{(p)} = \dot{K}_{\vartheta_1}^{(p)}$$

which means $K_0 = K_1$.

This completes the proof of our next theorem.

THEOREM 2. *The family* $\{\dot{K}_\vartheta^{(p)}\}$ *depends continuously on* (p, ϑ, K_0, K_1) *for* $1 \leqq p \leqq P < \infty$, $0 \leqq \vartheta \leqq 1$, K_i in $\mathscr{K}_1$. *It is strictly monotonic decreasing in* p *for* $1 \leqq p \leqq \infty$ *and convex in* ϑ.

An immediate consequence of Theorem 2 is as follows. Let $W_{(s)}(K)$ denote the sth cross-sectional measure of K, that is, the mixed volume

$$V(\underbrace{K, \ldots, K}_{(n-s)}; \underbrace{E, \ldots, E}_{s})$$

for $s = 0, 1, \ldots, n-1$. The measures $W_{(s)}(K)$ are well known to be monotonic in K, that is if $K \subseteq K'$ then $W_{(s)}(K) \leqq W_{(s)}(K')$ (cf. **(1)**, p. 50). Hence

$$W_{(s)}(\dot{K}_\vartheta^{(p)}) \geqq W_{(s)}(\dot{K}_\vartheta^{(q)})$$

when $1 \leqq p \leqq q \leqq \infty$, with equality if and only if K_0 and K_1 are identical. Thus $W_{(s)}(\dot{K}_\vartheta^{(p)})$ is monotonic decreasing in p and, in virtue of Theorem 2, continuous in that parameter. In particular, the intersection $K_0 \cap K_1$ has minimal cross-sectional measures and $\dot{K}_\vartheta^{(p)}$ has maximal. This latter family of bodies might well be called the set of weighted harmonic means of K_0 and K_1 in view of the next remarks.

A special instance of the convexity of the family $\dot{K}_\vartheta^{(1)}$ is

$$\dot{K}_\vartheta^{(1)} = [(1-\vartheta)\hat{K}_0 + \vartheta\hat{K}_1]\hat{} \subseteq (1-\vartheta)K_0 + \vartheta K_1.$$

In the inclusion, there is equality if and only if $K_0 = K_1$. This may be viewed as the analogue, for convex bodies, of the theorem of the arithmetic and harmonic means for positive numbers. Indeed, the latter may be looked upon as a special case of the former in which K_0 and K_1 are centrally symmetric bodies in a one-dimensional Euclidean space, the centre of symmetry being the common interior point Q. A similar observation is valid regarding the monotonicity of the means $\dot{K}_\vartheta^{(p)}$ in p for fixed ϑ.

The results of these last two paragraphs give us the inequalities

$$W_{(s)}(\dot{K}_\vartheta^{(p)}) \leqq W_{(s)}((1-\vartheta)K_0 + \vartheta K_1)$$

for $1 \leqq p \leqq \infty$ with equality if and only if $K_0 = K_1$. The next section furnishes an improvement on this result for the case $s = 0$, that is for the volume functional.

4. A dual Brunn–Minkowski theorem. For fixed p satisfying $1 \leqq p < \infty$, let $V(\dot{K}_\vartheta^{(p)}) = V_\vartheta$ be the volume of $\dot{K}_\vartheta^{(p)}$ where $0 \leqq \vartheta \leqq 1$. Since $\dot{K}_\vartheta^{(p)}$ contains an interior point Q, $V_\vartheta > 0$. The distance function of $\dot{K}_\vartheta^{(p)}$ is

$$F_\vartheta(x) = \sqrt[p]{[(1-\vartheta)F_0^p(x) + \vartheta F_1^p(x)]}.$$

Let

$$\bar{K}_i = \frac{1}{V_i^{1/n}} K_i;\ V(\bar{K}_i) = 1.$$

Set

$$\bar{F}_{\vartheta'}(x) = \sqrt[p]{[(1-\vartheta')\bar{F}_0^p(x) + \vartheta'\bar{F}_1^p(x)]}$$

where $\bar{F}_i(x) = V_i^{1/n}F_i(x)$ is the distance function of $\bar{K}_i$. Finally, let $\bar{V}_{\vartheta'}$ be the volume of that convex body whose distance function is $\bar{F}_{\vartheta'}(x)$. Since $F_\vartheta(x) = \bar{F}_{\vartheta'}(x)/\mu$, where

$$\mu = 1 \Big/ \sqrt[p]{\left[\frac{(1-\vartheta)}{V_0^{p/n}} + \frac{\vartheta}{V_1^{p/n}}\right]}, \vartheta' = \vartheta\mu^p/V_1^{p/n},$$

we have $V_\vartheta^{1/n} = \mu\bar{V}_{\vartheta'}^{1/n}$.

The polar co-ordinate formula for the volume of a convex body gives

$$\bar{V}_{\vartheta'} = \frac{1}{n}\int_{\partial E}\left[\frac{1}{\bar{F}_{\vartheta'}(u)}\right]^n dw,$$

where dw is the differential of surface area of the unit sphere E centred at Q.

For the integrand we have

$$1\Big/\sqrt[p]{\left[\frac{(1-\vartheta')}{\left(\frac{1}{\bar{F}_0(u)}\right)^p}+\frac{\vartheta'}{\left(\frac{1}{\bar{F}_1(u)}\right)^p}\right]} \leqq \sqrt[n]{\left[(1-\vartheta')\left(\frac{1}{\bar{F}_0(u)}\right)^n+\vartheta'\left(\frac{1}{\bar{F}_1(u)}\right)^n\right]}$$

with equality if and only if $\bar{F}_0(u) = \bar{F}_1(u)$. Therefore

$$\bar{V}_{\vartheta'} \leqq \frac{1}{n}\int_{\partial E}\left[\frac{(1-\vartheta')}{(\bar{F}_0(u))^n}+\frac{\vartheta'}{(\bar{F}_1(u))^n}\right]dw = (1-\vartheta')V(\bar{K}_0)+\vartheta' V(\bar{K}_1) = 1.$$

There is equality if and only if $\bar{K}_0 = \bar{K}_1$. This gives as the analogue of the Brunn–Minkowski theorem: $V_\vartheta{}^{1/n} \leqq \mu$. There is equality if and only if $K_0 = \lambda K_1$, $\lambda = (V_0/V_1)^{1/n}$, the centre of homothety being at Q.

If $p = \infty$, we have $K_0 \cap K_1 \subseteq K_i$ and so $V(K_0 \cap K_1) \leqq \min(V_0, V_1)$. Clearly there is equality if and only if one of the bodies K_i is a subset of the other. The volume functional is monotonic under set inclusion and so, by Theorem 2, $V(K_0 \cap K_1) \leqq V(\dot{K}_\vartheta{}^{(p)})$ for $1 \leqq p < \infty$ with equality if and only if $K_0 = K_1$.

We collect these results in our last theorem.

THEOREM 3.

$$V^{1/n}(K_0 \cap K_1) \leqq V^{1/n}(\dot{K}_\vartheta^{(p)}) \leqq 1\Big/\sqrt[p]{\left[\frac{(1-\vartheta)}{V^{p/n}(K_0)}+\frac{\vartheta}{V^{p/n}(K_1)}\right]},$$

for $1 \leqq p < \infty$. There is equality on the left if and only if $K_0 = K_1$ and on the right if and only if $K_0 = \lambda K_1$ with centre of homothety at Q. Further

$$V^{1/n}(K_0 \cap K_1) = V^{1/n}(K_\vartheta^{(\infty)}) \leqq \min(V^{1/n}(K_0), V^{1/n}(K_1))$$

with equality on the right if and only if $K_0 = K_1$.

REFERENCES

1. T. Bonnesen and W. Fenchel, *Theorie der Konvexen Körper* (New York, 1948).
2. W. Firey, *p-Means of convex bodies*, submitted to Math. Scand. (Oct., 1959).

Washington State University

ON THE HAUSDORFF AND TRIGONOMETRIC MOMENT PROBLEMS

P. G. ROONEY

Let K be a subset of $BV(0, 1)$—the space of functions of bounded variation on the closed interval $[0, 1]$. By the Hausdorff moment problem for K we shall mean the determination of necessary and sufficient conditions that corresponding to a given sequence $\mu = \{\mu_n | n = 0, 1, 2, \ldots\}$[1] there should be a function $\alpha \in K$ so that

(1) $$\mu_n = \int_0^1 t^n d\alpha(t), \quad n = 0, 1, 2, \ldots .$$

For various collections K this problem has been solved—see (**3**, Chapter III).

By the trigonometric moment problem for K we shall mean the determination of necessary and sufficient conditions that corresponding to a sequence $c = \{c_n | n = 0, \pm 1, \pm 2, \ldots\}$[2] there should be a function $\alpha \in K$ so that

(2) $$c_n = \int_0^1 e^{-2n\pi it} d\alpha(t), \quad n = 0, \pm 1, \pm 2, \ldots .$$

For various collections K this problem has also been solved—see, for example (**4**, Chapter IV, § **4**). It is noteworthy that these two problems have been solved for essentially the same collections K.

Recently (**2**), we gave new solutions of the trigonometric moment problem for certain classes K, namely those K determined by $K' = L_p(0, 1)$, $1 < p \leqslant 2$, where K' is defined now and henceforth, if the functions of K are absolutely continuous, to consist of all functions equal almost everywhere to the derivative of a function in K. These solutions were determined by use of the known solutions of the Hausdorff moment problem for these particular classes K.

Here we propose to generalize this procedure. Specifically, we propose to show that if the Hausdorff moment problem can be solved for a particular class K, then so can the trigonometric moment problem. This forms the content of the theorem below, and we shall illustrate our theory in a number of cases.

Received June 6, 1960.

[1]We shall use μ as a generic symbol for sequences whose indices run from zero to infinity.

[2]We shall use c as a generic symbol for sequences whose indices run from minus infinity to infinity.

To this end, we must first establish a number of results concerning certain numbers $a_{r,m}$ defined by

(3) $$a_{r,m} = \int_0^1 t^m e^{2\pi i r t} dt, \; r = 0, +1, \pm 2, \dots, \quad m = 0, 1, 2, \dots .$$

Since these numbers are essentially both the Hausdorff moments of the trigonometric powers of t, and the trigonometric moments of the algebraic powers of t, it is perhaps not surprising that they have an important role to play. Their properties are given in the following lemmas.

LEMMA 1.

(4) $$a_{r,m} = (1 - m a_{r,m-1})/2\pi i r, \qquad rm \neq 0,$$

(5) $$|a_{r,m}| < (m+1)^{-1},$$

(6) $$|a_{r,m}| < (\pi|r|)^{-1}, \qquad r \neq 0,$$

(7) $$a_{r,m} = \begin{cases} (m+1)^{-1}, r = 0, \\ 0, \quad m = 0, r \neq 0, \\ \sum_{n=0}^{m-1} \binom{m}{n} (-1)^n n!/(2\pi i r)^{n+1}, rm \neq 0. \end{cases}$$

Proof. On integration by parts, (4) follows from (3). If $m \neq 0$, (6) comes from applying (5), which is trivial, to the right-hand side of (4). The first two parts of (7) are immediate, and the third part follows from the second on repeated application of (4). From (7), (6) is obvious if $m = 0$.

LEMMA 2. *If* $|c_r| < M$, $r = 0, \pm 1, \pm 2, \dots$, *and*

$$\lim_{N\to\infty} \sum_{-N}^{N}{}' \frac{c_r}{r}$$

exists, (where the prime denotes the omission of the term corresponding to $r = 0$)*, then for* $m = 0, 1, 2, \dots$

$$\lim_{N\to\infty} \sum_{r=-N}^{N} c_r a_{r,m}$$

exists.

Proof. Since, from (7), $a_{r,0} = 0$, $r \neq 0$, and $a_{0,0} = 1$, it follows that

(8) $$\sum_{-N}^{N} c_r a_{r,0} = c_0,$$

and the limit exists for $m = 0$. Now if $m > 0$, then from (7) and (4),

(9) $$\sum_{r=-N}^{N} c_r a_{r,m} = \frac{1}{m+1} c_0 + \frac{1}{2\pi i} \sum_{-N}^{N}{}' \frac{c_r}{r} - \frac{m}{2\pi i} \sum_{r=-N}^{N}{}' \frac{c_r}{r} a_{r,m-1}.$$

But the first two terms on the right of this equation have limits as $N \to \infty$, so that it suffices to show that the third term has such a limit. But from (6)

$$\sum_{r=-\infty}^{\infty}{}' \left| \frac{c_r}{r} a_{r,m-1} \right| < \frac{M}{\pi} \sum_{-\infty}^{\infty}{}' \frac{1}{r^2} = \frac{\pi M}{3}, \tag{10}$$

so that the series

$$\sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1}$$

converges absolutely. Thus the limit of the last term in (9) also exists, and the lemma is proved.

With each sequence c, satisfying the hypotheses of Lemma 2 we can now associate a sequence $\mu(c)$ defined by

$$\mu_m(c) = \lim_{N\to\infty} \sum_{r=-N}^{N} c_r a_{r,m}. \tag{11}$$

The sequence $\mu(c)$ has certain properties that we summarize as a lemma.

LEMMA 3. *If c satisfies the hypotheses of Lemma 2, then*

$$\mu_0(c) = c_0, \tag{12}$$

$$\mu_m(c) = \frac{1-m}{1+m} \frac{c_0}{2} + \mu_1 - \frac{m}{2\pi i} \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1}\ m > 0. \tag{13}$$

Proof. Equation (12) follows immediately from (8) and (11). Now from (7), $a_{r,1} = (2\pi i r)^{-1}$, $r \neq 0$, $a_{0,1} = \frac{1}{2}$. Hence from (11) and (9),

$$\begin{aligned}
\mu_m(c) &= \lim_{N\to\infty} \sum_{r=-N}^{N} c_r a_{r,m} \\
&= \lim_{N\to\infty} \left(\frac{1}{m+1} c_0 + \frac{1}{2\pi i} \sum_{-N}^{N}{}' \frac{c_r}{r} - \frac{m}{2\pi i} \sum_{-N}^{N}{}' \frac{c_r}{r} a_{r,m-1} \right) \\
&= \lim_{N\to\infty} \left(\frac{1}{m+1} c_0 - \frac{1}{2} c_0 + \sum_{-N}^{N} c_r a_{r,1} - \frac{m}{2\pi i} \sum_{-N}^{N}{}' \frac{c_r}{r} a_{r,m-1} \right) \\
&= \frac{1-m}{1+m} \frac{c_0}{2} + \mu_1 - \frac{m}{2\pi i} \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1},
\end{aligned}$$

since by (10), this last series converges absolutely.

We are now ready to state and prove our theorem.

THEOREM. *Necessary and sufficient conditions that a sequence c be represented in the form* (2) *for some $\alpha \in K$ are that*

(i) $|c_r| < M$, $r = 0, \pm 1, \pm 2, \ldots,$

(ii) $\lim_{N\to\infty} \sum_{-N}^{N}{}' \frac{c_r}{r}$ *exists,*

(iii) *$\mu(c)$ is represented in the form* (1) *with $\alpha \in K$.*

Proof of necessity. Suppose

$$c_n = \int_0^1 e^{-2n\pi i t}\, d\alpha(t),\ n = 0, \pm 1, \pm 2, \dots ,$$

where $\alpha \in K$. Clearly,

$$|c_n| < \int_0^1 dV(t) = M,$$

where $V(t)$ is the total variation of α, so that (i) is necessary.

Now let

$$\beta(t) = 2\pi\alpha(t/2\pi) - c_0 t, \qquad 0 < t < 2\pi,$$

and define $\beta(t)$ outside this interval by

$$\beta(t + 2\pi) = \beta(t).$$

Then $\beta(t)$ is periodic and of bounded variation, so that if

$$d_n = \frac{1}{2\pi}\int_0^{2\pi} e^{-int}\beta(t)dt,\ n = 0, \pm 1, \pm 2, \dots ,$$

it follows from the Dini–Dirichlet test (**4**, Chapter II, Theorem 8.1)

$$\lim_{N\to\infty} \sum_{-N}^{N} d_n = \frac{1}{2}\,(\beta(0+) + \beta(0-)). \tag{14}$$

But if $n \neq 0$,

$$d_n = \frac{1}{2\pi}\int_0^{2\pi} e^{-int}\beta(t)dt = \int_0^{2\pi} e^{-int}\alpha(t/2\pi)dt - \frac{c_0}{2\pi}\int_0^{2\pi} te^{-int}dt$$

$$= 2\pi\int_0^1 e^{-2n\pi i t}\alpha(t)dt + \frac{c_0}{in},$$

and integrating by parts, we obtain, if $n \neq 0$,

$$d_n = c_n/in.$$

Thus, (14) becomes

$$\lim_{N\to\infty} \sum_{-N}^{N}{}' \frac{c_r}{r} = \frac{i}{2}\,(\beta(0+) + \beta(0-)) - i\, d_0,$$

and (ii) is necessary.

Now let

$$d_n' = \frac{1}{2\pi}\int_0^{2\pi} e^{-int}d\beta(t).$$

Then, since

$$a_{r,n} = \int_0^1 t^n\, e^{2\pi irt}dt = (2\pi)^{-(n+1)}\int_0^{2\pi} t^n\, e^{irt}dt,$$

it follows from Parseval's theorem for Fourier series (**4**, Chapter IV, Theorem 8.7 (iv)) that

$$\int_0^{2\pi} t^n\, d\beta(t) = \sum_{r=-\infty}^{\infty} d'_r (2\pi)^{n+1} a_{r,n}\ (C, 1). \tag{15}$$

But the left-hand side of (15) is equal to

$$\int_0^{2\pi} t^n d\beta(t) = 2\pi \int_0^{2\pi} t^n d\alpha(t/2\pi) - c_0 \int_0^{2\pi} t^n dt$$

$$= (2\pi)^{n+1} \left\{ \int_0^1 t^n d\alpha(t) - c_0/(n+1) \right\}.$$

Also

$$d'_n = \frac{1}{2\pi}\int_0^{2\pi} e^{-int} d\beta(t) = \int_0^{2\pi} e^{-int} d\alpha(t/2\pi) - \frac{c_0}{2\pi}\int_0^{2\pi} e^{-int} dt$$

$$= \int_0^1 e^{-2n\pi it} d\alpha(t) - \frac{c_0}{2\pi}\int_0^{2\pi} e^{-int} dt$$

$$= \begin{cases} c_n & n \neq 0 \\ 0 & n = 0, \end{cases}$$

so that (15) becomes

$$(2\pi)^{n+1} \left\{ \int_0^1 t^n d\alpha(t) - \frac{c_0}{n+1} \right\} = (2\pi)^{n+1} {\sum_{-\infty}^{\infty}}' c_r a_{r,n} \qquad (C, 1).$$

Thus, since $a_{0,n} = (n+1)^{-1}$

$$\int_0^1 t^n d\alpha(t) = \sum_{-\infty}^{\infty} c_r a_{r,n} \qquad (C, 1).$$

But, since by (i), (ii), and Lemma 2

$$\lim_{N\to\infty} \sum_{r=-N}^{N} c_r a_{r,n}$$

exists and equals $\mu_n(c)$, and since the $(C, 1)$ method is consistent, we must have

$$\mu_n(c) = \int_0^1 t^n d\alpha(t),$$

and (iii) is necessary.

Proof of sufficiency. From (iii), $\alpha \in K$ exists so that

$$\mu_n(c) = \int_0^1 t^n d\alpha(t).$$

Let

$$c'_n = \int_0^1 e^{-2n\pi it} d\alpha(t).$$

We shall show that $c_n' = c_n$. Firstly, $c_0' = c_0$, for from (12),

$$c_0 = \mu_0(c) = \int_0^1 d\alpha(t) = c_0'.$$

Then, if $n \neq 0$,

$$\begin{aligned} c_n' &= \int_0^1 e^{-2n\pi i t} d\alpha(t) \\ &= \sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} \int_0^1 t^m d\alpha(t) = \sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} \mu_m(c), \end{aligned}$$

the interchange of integration and summation being justified by the uniform convergence of the exponential series.

But then using (13) and (12), if $n \neq 0$,

$$\begin{aligned} (16) \quad c_n' &= \sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} \mu_m(c) \\ &= \mu_0(c) + \sum_{m=1}^{\infty} \frac{(-2n\pi i)^m}{m!} \left(\frac{1-m}{1+m} \cdot \frac{c_0}{2} + \mu_1 - \frac{m}{2\pi i} \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1} \right) \\ &= c_0 \left(1 + \frac{1}{2} \sum_{m=1}^{\infty} \frac{(1-m)(-2n\pi i)^m}{(m+1)!} \right) + \mu_1 \sum_{m=1}^{\infty} \frac{(-2n\pi i)^m}{m!} \\ &\qquad + n \sum_{m=1}^{\infty} \frac{(-2n\pi i)^{m-1}}{(m-1)!} \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1}. \end{aligned}$$

Now

$$\sum_{m=1}^{\infty} \frac{x^m}{(m+1)!} = \frac{e^x - 1}{x} - 1,$$

and

$$\sum_{m=1}^{\infty} \frac{mx^m}{(m+1)!} = e^x - \frac{e^x - 1}{x},$$

so that the coefficient of c_0 in (16) is equal to

$$1 + \frac{1}{2} \left(\frac{e^{-2n\pi i} - 1}{-2n\pi i} - 1 - e^{-2n\pi i} + \frac{e^{-2n\pi i} - 1}{-2n\pi i} \right) = 0.$$

Also,

$$\sum_{m=1}^{\infty} \frac{x^m}{m!} = e^x - 1,$$

so that the coefficient of μ_1 in (16) is also zero. Thus if $n \neq 0$

$$(17) \qquad c_n' = n \sum_{m=1}^{\infty} \frac{(-2n\pi i)^{m-1}}{(m-1)!} \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} a_{r,m-1}.$$

But the series

$$\sum_{m=1}^{\infty} \frac{|(-2n\pi i)|^{m-1}}{(m-1)!} \sum_{-\infty}^{\infty}{}' \left|\frac{c_r}{r} a_{r,m-1}\right| < \infty.$$

For from (10) it is smaller than

$$\frac{\pi M}{3} \sum_{m=1}^{\infty} \frac{(2|n|\pi)^{m-1}}{(m-1)!} = \frac{\pi M}{3} e^{2|n|\pi} < \infty.$$

Hence we can interchange the orders of summation in (17) and obtain

$$c_n' = n \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} \sum_{m=1}^{\infty} \frac{(-2n\pi i)^{m-1}}{(m-1)!} a_{r,m-1} \tag{18}$$

$$= n \sum_{-\infty}^{\infty}{}' \frac{c_r}{r} \sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} a_{r,m}.$$

But

$$\sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} a_{r,m} = \sum_{m=0}^{\infty} \frac{(-2n\pi i)^m}{m!} \int_0^1 t^m e^{2r\pi i t} dt$$

$$= \int_0^1 \left(\sum_{m=0}^{\infty} \frac{(-2n\pi i t)^m}{m!} \right) e^{2r\pi i t} dt = \int_0^1 e^{2(r-n)\pi i t} dt$$

$$= \begin{cases} 0, & r \neq n \\ 1, & r = n, \end{cases}$$

and using this in (18) we obtain $c_n' = c_n$, that is

$$c_n = \int_0^1 e^{-2n\pi i t} d\alpha(t), \quad n = 0, \pm 1, \pm 2, \dots ,$$

with $\alpha \in K$.

As an example of the use of the theorem to obtain solutions of the trigonometric moment problem, let us take $K = BV(0, 1)$. Then from (**3**, Chapter III, Theorem 2b), a necessary and sufficient condition that sequence μ be the Hausdorff moment sequence of a function in $BV(0, 1)$ is that for some constant L

$$\sum_{m=0}^{k} |\lambda_{k,m}| < L, \quad k = 0, 1, 2, \dots ,$$

where

$$\partial_{j,m} = \binom{k}{m} (-1)^{k-m} \Delta^{k-m} \mu_m,$$

and Δ is the advancing difference operator.

Thus, given a sequence c, we find as necessary and sufficient conditions that c be the trigonometric moment sequence of a function in $BV(0, 1)$, are that (i) and (ii) of the theorem be satisfied, and that for some constant L,

$$\sum_{m=0}^{k} |\lambda_{k,m}(c)| < L, \; k = 0, 1, 2, \ldots ,$$

where

$$\lambda_{k,m}(c) = \binom{k}{m} (-1)^{k-m} \Delta^{k-m} \mu_m \tag{19}$$

$$= \lim_{N\to\infty} \sum_{r=-N}^{N} c_r \, a_{r,k,m},$$

where

$$a_{r,k,m} = \binom{k}{m} (-1)^{k-m} \Delta^{k-m} a_{r,m} = \binom{k}{m} \int_0^1 t^m \, (1-t)^{k-m} e^{2\pi i r t} dt.$$

We list in Table I the conditions for representation as a trigonometric moment sequence for some of the more common classes K. In all cases (i) and (ii) of the theorem must hold and the column marked (iii) gives the third condition that must hold. The last column gives the place from which the conditions for the Hausdorff representation are taken.

TABLE I

	K	(iii)	Reference (3, Chapter III)
1	$BV\,(0,1)$	$\sum_{m=0}^{k} \lvert\lambda_{k,m}(c)\rvert < L,\ k = 0, 1, 2, \ldots.$	Theorem 2b
2	Increasing functions on [0,1]	$\lambda_{k,m}(c) > 0,\ k = 0, 1, 2, \ldots, 0 < m < k,$	Theorem 4a
3	$K' = L_p(0, 1), 1 < p < \infty$	$(k+1)^{p-1} \sum_{m=0}^{k} \lvert\lambda_{k,m}(c)\rvert^p < L,\ k = 0, 1, 2, \ldots.$	Theorem 5
4	$K' = L_\infty\,(0, 1)$	$(k+1)\lvert\lambda_{r,m}\rvert < L,\ k = 0, 1, 2, \ldots,\ 0 < m < k.$	Theorem 6

Case 2 is of particular note, since the trigonometric moment problem for this K was given a particularly elegant solution by Bochner (1, § 20).

References

1. S. Bochner, *Vorlesungen über Fouriersche Integrale* (Leipzig, 1932).
2. P. G. Rooney, *On the representation of sequences as Fourier coefficients*, Proc. Amer. Math. Soc., *11* (1960), 762–768.
3. D. V. Widder, *The Laplace transform* (Princeton, 1941).
4. A. Zygmund, *Trigonometric series I* (Cambridge, 1959).

University of Toronto

THE EXPANSION PROBLEM WITH BOUNDARY CONDITIONS AT A FINITE SET OF POINTS

RANDAL H. COLE

1. Introduction. The problem of expanding an arbitrary function in a series of characteristic solutions of the ordinary differential equation

$$u^{[n]} + P_1 u^{[n-1]} + \ldots + P_n u = 0 \qquad \left(u^{[j]} = \frac{d^j u}{dx^j}\right) \tag{1.1}$$

and the boundary relations

$$\sum_{\mu=1}^{m} \sum_{j=1}^{n} v_{ij}^{(\mu)} u^{[j-1]}(a_\mu) = 0, \qquad i = 1, 2, \ldots, n, \tag{1.2}$$

is well known. The various discussions are distinguished by the manner in which a parameter λ appears in the differential system and by the number of points at which the boundary conditions apply. The case in which the boundary conditions apply at intermediate as well as at the end points of a fundamental interval has been considered by Wilder **(3)**. His investigation was confined to the case where $P_n = P_{n0}(x) + \lambda^n$ and where each coefficient $v_{ij}^{(\mu)}$ in the boundary relations is free from λ.

The present discussion treats the case where each coefficient P_k is a polynomial in λ of degree k and each coefficient $v_{ij}^{(\mu)}$ in the boundary relations is an arbitrary polynomial in λ. The reduction of the system (1.1) and (1.2) to an equivalent matrix system has been accomplished **(4)**, therefore the results obtained by Langer **(1)** can be applied to the present problem.[1] It will be assumed that the reader is familiar with Langer's paper so that direct reference can be made to some of his formulas. In order to facilitate the use of such formulas, Langer's notation has been used here with only minor modifications.

Langer's development concerns a differential system in the complex domain. His boundary conditions apply at a specified set of m points in this domain. Although his results are valid when the variable is restricted to be real, there are several points of interest attending this restriction. The first of these is the form of Green's matrix. Langer has defined a set of m Green's matrices corresponding to the m boundary points. In the real case, these can be combined to yield a single Green's matrix, $\mathfrak{G}(x, s, \lambda)$, which has a finite discontinuity, with respect to the variable s, at each of the boundary points. In all

Received June 23, 1960.

[1]The author wishes to thank Professor R. E. Langer for valuable suggestions. Some of the results contained in this paper were obtained at the 1958 Summer Research Institute of the Canadian Mathematical Congress.

other respects, this matrix has the familiar properties of a Green's function. That is, it has a unit discontinuity when $x = s$, and it is a formal solution of the given boundary system and of the adjoint system. The second point of interest is that the adjoint boundary conditions (3·5b) are simply specifications of finite discontinuities at the boundary points. The discontinuities of the Green's matrix satisfy these adjoint conditions. It is clear, therefore, that these finite jumps are characteristic of the adjoint solution and of Green's matrix and, further, that no system with boundary conditions of the form of (2·1b) can be self-adjoint if $m > 2$.

It should be noted that Haltiner **(5)** specialized Langer's results to the real case for two point boundary conditions and obtained a new definition of adjoint boundary relations. These relations have the advantage of being explicitly defined in terms of the given boundary problem. The same advantage is enjoyed by the m point relations obtained here.

The formal points of interest outlined above are significant, but the primary problem in the subsequent discussion is the determination of the specific regularity conditions on the boundary problem which will ensure the convergence of the expansion of an arbitrary vector. This is accomplished by decomposing the Green's matrices defined by Langer and by finding relations among the parts. These relations are equally valid in the complex case and can be used to broaden the scope of Langer's regularity conditions. This point will be amplified in § 5, but it is appropriate to point out here that Langer's general results have been illuminated by applying them to a special case.

Whyburn **(6)**, **(7)**, **(8)** has considered differential systems in which integral boundary conditions are combined with linear conditions at a countable set of points. In particular **(6)**, he has developed some formal aspects of a system with combined integral and two point conditions. His Green's matrix is consistent with the Green's matrix defined below and will, therefore, lend itself to a reduction similar to that achieved in § 4.

2. The differential system. The basic system is the equation (1.1) and boundary conditions (1.2) with the following assumptions:

(a) $$P_k = \sum_{l=0}^{k} P_{kl}(x)\lambda^l, \quad k = 1, 2, \ldots, n,$$

with $P_{kl}(x)$ free from λ and indefinitely differentiable.

(b) *The algebraic equation $r^n + P_{11}(x)r^{n-1} + \ldots + P_{nn}(x) = 0$ has roots $r_i(x)$, $i = 1, 2, \ldots, n$, which together with their differences, $r_i(x) - r_j(x)$, $i \neq j$, have constant arguments and are bounded from zero for all values of x on a fundamental interval $[a_1, a_m]$.*

(c) *The points $a_1, a_2, \ldots, a_m$, $(a_i < a_{i+1})$, at which the boundary relations apply, are the end points of the fundamental interval and a set of $m - 2$ arbitrary interior points of that interval.*

(d) *Each coefficient $v_{ij}^{(\mu)}$ in (1.2) is an arbitrary polynomial in λ with constant coefficients.*

This system can be reduced to the matrix system (see **(4)**)

$$\mathfrak{Y}'(x, \lambda) = \{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{Y}(x, \lambda) \tag{2.1a}$$

$$\sum_{\mu=1}^{m} \mathfrak{W}^{(\mu)}(\lambda)\mathfrak{Y}(a_\mu, \lambda) = \mathfrak{O}, \tag{2.1b}$$

where $\mathfrak{R}(x)$ is the diagonal matrix $(\delta_{ij}r_j(x))$; the diagonal components of the matrix $\mathfrak{Q}(x)$ are zeros, and the other components are indefinitely differentiable and free from λ; and the components of $\mathfrak{W}^{(\mu)}(\lambda)$ are polynomials in λ.

The above results may be stated as a theorem.

THEOREM 1. *The system* (1.1) *and* (1.2), *satisfying assumptions* (a), (b), (c), *and* (d), *may be reduced to the matrix system* (2.1a) *and* (2.1b).

All the subsequent results are developed for the matrix system which, therefore, can well be regarded as the basic one. The *n*th order system is preferred in this role because of its classical significance.

Langer has treated the problem associated with the matrix system when x is a complex variable and the boundary points are m specified points in the x-plane. He has obtained asymptotic solutions of the equation, defined the adjoint system and a set of m Green's matrices, developed a biorthogonality relation and a formal expansion of an arbitrary vector in a series of characteristic solutions. He has expressed the expansion as a series of residues of Green's matrices and shown that under appropriate conditions the latter converges to the arbitrary vector. Langer's formal results will be adapted to the present problem. An independent derivation of Green's matrix and of the formal expansion would contribute to the continuity of this discussion but would to some extent duplicate known results. Furthermore, such a derivation can be applied to more general boundary conditions than those considered here and will be made the subject of a separate discussion. The pertinent results from Langer's paper are given below, some of them being stated in the form of theorems.

The characteristic values, $\lambda_1, \lambda_2, \ldots$, of system (2.1) are the roots of the equation $D(\lambda) = 0$ (cf. **(1**, § 7**)**). $D(\lambda)$ is the determinant of the matrix

$$\mathfrak{D}(\lambda) = \sum_{\mu=1}^{m} \mathfrak{W}^{(\mu)}(\lambda)\mathfrak{Y}(a_\mu, \lambda) \tag{2.2}$$

where $\mathfrak{Y}(x, \lambda)$ is any non-singular matrix solution of (2.1a). The characteristic solutions are non-trivial vector solutions of (2.1). They exist when λ is a characteristic value. The Green's matrices, $\mathfrak{G}^{(\mu)}(x, s, \lambda)$, $\mu = 1, 2, \ldots, m$, are defined by (see **(1**, § 9**)**)

$$\mathfrak{G}^{(\mu)}(x, s, \lambda) = \mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda)\mathfrak{W}^{(\mu)}(\lambda)\mathfrak{Y}(a_\mu, \lambda)\mathfrak{Y}^{-1}(s, \lambda), \tag{2.3}$$

where $\mathfrak{Y}(x, \lambda)$ is the non-singular matrix solution used in the definition of $\mathfrak{D}(\lambda)$. Let τ be a non-negative integer and let $\mathfrak{f}(x)$ be any vector (n-tuple of

real functions) which has a derivative of order $\tau + 1$. Define the set of vectors, $\mathfrak{f}^{(0)}(x), \mathfrak{f}^{(1)}(x), \ldots, \mathfrak{f}^{(\tau+1)}(x)$, by the relations **(1**, (15.3)**)**

$$\mathfrak{f}^{(0)}(x) = \mathfrak{f}(x),$$
$$\mathfrak{f}^{(h)}(x) = \mathfrak{R}^{-1}(x)\{\mathfrak{f}^{(h-1)\prime}(x) - \mathfrak{Q}(x)\mathfrak{f}^{(h-1)}(x)\}, \quad h = 1, 2, \ldots, \tau + 1.$$

THEOREM 2 (*cf.* **(1**, (15.8)**)**). *The formal expansion of* $\mathfrak{f}^{(l)}(x)$ *may be reduced to the infinite series of residues*

$$\mathfrak{s}^{(l)}(x) = \sum_{\beta=0}^{\infty} \operatorname{res}_\beta \sum_{\mu=1}^{m} \lambda^l \left\{ \int_{a_1}^{a_\mu} \mathfrak{G}^{(\mu)}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds \right.$$
$$\left. + \mathfrak{G}^{(\mu)}(x, a_\mu, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1}\mathfrak{f}^{(h)}(a_\mu) \right\}.$$

THEOREM 3. *The partial sum,* $\mathfrak{s}_k{}^{(l)}(x)$, *of the series of residues associated with the first* k *characteristic values, is given by*

$$\mathfrak{s}_k^{(l)}(x) = \frac{1}{2\pi i}\int_{\Gamma_k} \sum_{\mu=1}^{m} \left\{ -\int_{a_\mu}^{x} \mathfrak{G}^{(\mu)}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds \right.$$
$$\left. + \mathfrak{G}^{(\mu)}(x, a_\mu, \lambda) \sum_{h=0}^{\tau} \lambda^{-h-1}\mathfrak{f}^{(h)}(a_\mu) \right\} \lambda^l d\lambda, \tag{2.4}$$

where Γ_k *is a contour in the* λ*-plane enclosing precisely the first* k *characteristic values.*

The relation (2.4) is Langer's formula **(1**, (15.10)**)** except that x_1 has been replaced by s and η_μ by a_μ. It is clear that the expansion depends on the choice of the integer τ and that if $l = 0$, we have an expansion of the vector $\mathfrak{f}(x)$ itself.

3. Green's matrix.

The term

$$\sum_{\mu=1}^{m} \int_{a_\mu}^{x} \mathfrak{G}^{(\mu)}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds,$$

appearing in formula (2.4), represents the sum of m integrals. In the complex case, each integral is over a curve joining one of the boundary points to the point x. These curves may be entirely distinct or they may be drawn so that they have segments in common. In the real case, on the other hand, no such option exists. The intervals of integration have, of necessity, points in common. Consequently, in the real case it is notationally convenient to define $G(x, s, \lambda)$, which will be called the *Green's matrix*, by the relation

$$\mathfrak{G}(x, s, \lambda) = \left\{ \begin{array}{ll} \sum_{\mu=1}^{q} \mathfrak{G}^{(\mu)}(x, s, \lambda), & s < x \\ -\sum_{\mu=q+1}^{m} \mathfrak{G}^{(\mu)}(x, s, \lambda), & s > x \end{array} \right\}, \ s \text{ on } (a_q, a_{q+1}). \tag{3.1}$$

With $\mathfrak{G}(x, s, \lambda)$ thus defined by a distinct formula on each of the subintervals into which (a_1, a_m) is subdivided by the point x and the intermediate boundary points, it is easily verified that

$$(3.2) \qquad \sum_{\mu=1}^{m} \int_{a_\mu}^{x} \mathfrak{G}^{(\mu)}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds = \int_{a_1}^{a_m} \mathfrak{G}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds.$$

Employing (3.1) and (2.3), the discontinuities of $\mathfrak{G}(x, s, \lambda)$ at the boundary points are seen to be such that

$$(3.3) \qquad \begin{aligned} \mathfrak{G}(x, a_h + 0, \lambda) - \mathfrak{G}(x, a_h - 0, \lambda) &= \mathfrak{G}^{(h)}(x, a_h, \lambda) \\ &= \mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda)\mathfrak{W}^{(h)}(\lambda) \end{aligned}$$

where, as a notational convenience, the symbols $G(x, a_1 - 0, \lambda)$ and $G(x, a_m + 0, \lambda)$ are used to represent the zero matrix. In terms of Green's matrix, then, formula (2.4) becomes

$$(3.4) \qquad \mathfrak{s}_k^{(l)}(x) = \frac{1}{2\pi i}\int_{\Gamma_k}\Big[-\int_{a_1}^{a_m} \mathfrak{G}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}(s)ds$$
$$+ \sum_{\mu=1}^{m} \{\mathfrak{G}(x, a_\mu + 0, \lambda) - \mathfrak{G}(x, a_\mu - 0, \lambda)\} \sum_{h=0}^{r} \lambda^{-h-1}\mathfrak{f}^{(h)}(a_\mu)\Big]\lambda^l d\lambda.$$

It is of interest to observe that the Green's matrix defined in (3.1) has all the familiar properties of a Green's function in classical boundary problems. Because of its form, each matrix $\mathfrak{G}^{(\mu)}(x, s, \lambda)$, regarded as a function of x, is a solution of equation (2.1a). Since, therefore, $\mathfrak{G}(x, s, \lambda)$ is a sum of such matrices, it is a formal solution of (2.1a). It fails to be a true solution because of a discontinuity at $x = s$. Further, it is easily verified that

$$\sum_{h=0}^{m} \mathfrak{W}^{(h)}(\lambda)\mathfrak{G}(a_h, s, \lambda) = \mathfrak{O}.$$

Thus, $\mathfrak{G}(x, s, \lambda)$ is a formal matrix solution of the boundary problem (2.1).

The boundary problem adjoint to (2.1) may be defined by

$$(3.5a) \qquad \mathfrak{Z}'(x, \lambda) = -\mathfrak{Z}(x, \lambda)\{\lambda\mathfrak{R}(x) + \mathfrak{Q}(x)\}$$

$$(3.5b) \qquad \mathfrak{Z}(a_h + 0, \lambda) - \mathfrak{Z}(a_h - 0, \lambda) = \mathfrak{A}(\lambda)\mathfrak{W}^{(h)}(\lambda), \qquad h = 1, 2, \ldots, m,$$

where for convenience the symbols $\mathfrak{Z}(a_1 - 0, \lambda)$ and $\mathfrak{Z}(a_m + 0, \lambda)$ are defined to represent the zero matrix. A matrix $\mathfrak{Z}(x, \lambda)$ is a solution of this system if it satisfies equation (3.5a) and if a parametric matrix $\mathfrak{A}(\lambda)$ exists such that (3.5b) is satisfied. Solutions of the adjoint system, therefore, have discontinuities at the boundary points. This definition of the adjoint system is consistent with Langer's definition (**1**, (10.1)) if η_0 is identified with a_1.

Because of its form, Green's matrix, regarded as a function of s, is seen to be a formal solution of (3.5a). Moreover, recalling (3.3), its discontinuities at the boundary points are evidently precisely those required by (3.5b) with $\mathfrak{Y}(x, \lambda)\mathfrak{D}^{-1}(\lambda)$ as the parametric matrix $\mathfrak{A}(\lambda)$. As in the earlier case, it fails to be a true solution of (3.5) because of an additional discontinuity at $s = x$.

The characteristics of Green's matrix will be listed in the form of a theorem.

THEOREM 4. *Green's matrix defined by* (3.1) *has the following properties:*

(i) *It is continuous in x and s except when $x = s$ and when $s = a_\mu$, $\mu = 1, 2, \ldots, m$. The discontinuity when $x = s$ is given by*

$$\mathfrak{G}(s + 0, s, \lambda) - \mathfrak{G}(s - 0, s, \lambda) = \mathfrak{J}.$$

(ii) *For each fixed s, it is a formal solution of the boundary system* (2.1).

(iii) *For each fixed x, it is a formal solution of the boundary system* (3.5).

The non-homogeneous boundary problem,

$$\mathfrak{y}'(x, \lambda) = \{\lambda \mathfrak{R}(x) + \mathfrak{Q}(x)\}\mathfrak{y}(x, \lambda) + \mathfrak{f}(x)$$

$$\sum_{h=1}^{m} \mathfrak{W}^{(h)}(\lambda)\mathfrak{y}(a_h, \lambda) = \mathfrak{o},$$

when λ is not a characteristic value, has a vector solution $\mathfrak{u}(x, \lambda)$ given by

$$\mathfrak{u}(x, \lambda) = \int_{a_1}^{a_m} \mathfrak{G}(x, s, \lambda)\mathfrak{f}(s)ds.$$

The corresponding non-homogeneous adjoint problem has the vector solution

$$\mathfrak{v}(x, \lambda) = -\int_{a_1}^{a_m} \mathfrak{f}(s)\mathfrak{G}(s, x, \lambda)ds,$$

with

$$\mathfrak{a}(\lambda) = -\int_{a_1}^{a_m} \mathfrak{f}(s)\mathfrak{Y}(s, \lambda)ds\mathfrak{D}^{-1}(\lambda)$$

as the parametric vector. The verification of these facts is straightforward.

A reduction of formula (3.4) can be achieved by using the fact that $\mathfrak{G}(x, s, \lambda)$ is a formal solution of the adjoint system. It is more convenient, however, to cite the reduction given by Langer in **(1**, § 17**)** which results in his formula (17.3) and to express this latter formula in terms of the matrix $\mathfrak{G}(x, s, \lambda)$. The result is

$$(3.6) \quad \mathfrak{S}_k^{(l)}(x) = \frac{1}{2\pi i}\int_{\Gamma_k} \sum_{h=0}^{r} \mathfrak{f}^{(h)}(x)\lambda^{-h+l-1}d\lambda - \frac{1}{2\pi i}\int_{\Gamma_k}\int_{a_1}^{a_m} \lambda^{l-r-1}\mathfrak{G}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}^{(r+1)}(s)dsd\lambda.$$

Since the first term on the right of (3.6) has the value $\mathfrak{f}^{(l)}(x)$, we may write

$$(3.7) \quad \mathfrak{S}_k^{(l)}(x) = \mathfrak{f}^{(l)}(x) - \frac{1}{2\pi i}\mathfrak{b}_k^{(l)}(x)$$

where

$$(3.8) \quad \mathfrak{b}_k^{(l)}(x) = \int_{\Gamma_k}\int_{a_1}^{a_m} \lambda^{l-r-1}\mathfrak{G}(x, s, \lambda)\mathfrak{R}(s)\mathfrak{f}^{(r+1)}(s)dsd\lambda.$$

Thus the problem of showing that $\mathfrak{s}_k^{(l)}(x)$ converges to $\mathfrak{f}^{(l)}(x)$ has been reduced to the problem of showing that

$$\lim_{k\to\infty} \mathfrak{d}_k^{(l)}(x) = \mathfrak{o}.$$

4. The structure of Green's matrix. The synthesis of Green's matrix, achieved by formula (3.1), is notationally convenient in dealing with the formal aspects of the boundary problem. In order to establish the convergence of the formal expansion, however, it is desirable to obtain a decomposition of Green's matrix beyond that exhibited in (3.1). The following lemma will be useful in this reduction.

LEMMA 1. *Let $\mathfrak{U}^{(1)}, \mathfrak{U}^{(2)}, \ldots, \mathfrak{U}^{(m)}$ be a set of $n \times n$ matrices and let their sum, $\mathfrak{D}$, be non-singular. Corresponding to each matrix $\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)}$, there is a set of $(m+1)$ matrices $\mathfrak{H}^{(\mu\nu)}$, $\nu = 1, 2, \ldots, m+1$, such that*

$$\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)} = \sum_{\nu=1}^{m+1} \mathfrak{H}^{(\mu\nu)}, \qquad \mu = 1, 2, \ldots, m,$$

and

$$\mathfrak{H}^{(\mu\nu)} = -\,\mathfrak{H}^{(\nu\mu)}, \qquad \mu, \nu = 1, 2, \ldots, m.$$

The matrix $\mathfrak{H}^{(\mu, m+1)}$ has zero components except on its diagonal where each component is the corresponding diagonal component of $\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)}$.

Let the symbol $\mathfrak{J}_{hl}$ represent the matrix in which all the components are zero except for a unit component in the hth row and lth column. That is, $\mathfrak{J}_{hl} = (\delta_{ih}\delta_{lj})$, $h, l = 1, 2, \ldots, n$. Also, let the matrix $\mathfrak{J}^{hh}$ be defined by $\mathfrak{J}^{hh} = \mathfrak{J} - \mathfrak{J}_{hh}$. The cofactor of the element in the jth row and the ith column of $\mathfrak{D}$ may be written as $|\mathfrak{D}\mathfrak{J}^{ii} + \mathfrak{J}_{ji}|$. Hence, if D is the determinant of $\mathfrak{D}$,

$$\mathfrak{D}^{-1} = 1/D\,(|\mathfrak{D}\mathfrak{J}^{ii} + \mathfrak{J}_{ji}|)$$

and

$$\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)} = 1/D\left(\sum_{k=1}^{n} |\mathfrak{D}\mathfrak{J}^{ii} + \mathfrak{J}_{ki}|u_{kj}^{(\mu)}\right)$$

$$= 1/D\left(\sum_{k=1}^{n} |\mathfrak{D}\mathfrak{J}^{ii} + u_{kj}^{(\mu)}\mathfrak{J}_{ki}|\right).$$

The general component of the matrix on the right is exhibited as the sum of n determinants which differ from each other only with respect to their ith columns. They may be added, therefore, by replacing the ith column of any one of them by the sum of the ith columns of all of them. Since this column sum is readily seen to be the jth column of $\mathfrak{U}^{(\mu)}$, we have

$$\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)} = 1/D(|\mathfrak{D}\mathfrak{J}^{ii} + \mathfrak{U}^{(\mu)}\mathfrak{J}_{ji}|).$$

The right side of this relation may be expressed as the sum of two matrices, one having zeros on the diagonal and the other having zeros elsewhere. Thus,

$$(4.1)\quad \mathfrak{D}^{-1}\mathfrak{U}^{(\mu)} = \left(\frac{1-\delta_{ij}}{D}\,|\mathfrak{D}\mathfrak{I}^{ii} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}|\right) + \left(\frac{\delta_{ij}}{D}\,|\mathfrak{D}\mathfrak{I}^{ii} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}|\right).$$

The second matrix on the right will be represented by the symbol $\mathfrak{H}^{(\mu,m+1)}$. Since it is a diagonal matrix, we may replace the index i by j so that,

$$(4.2)\qquad \mathfrak{H}^{(\mu,m+1)} = \left(\frac{\delta_{ij}}{D}\,|\mathfrak{D}\mathfrak{I}^{jj} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{jj}|\right).$$

The first matrix on the right of (4.1) may be decomposed into a sum of m matrices by expanding the determinantal factor of the general component as follows.

$$\begin{aligned}|\mathfrak{D}\mathfrak{I}^{ii} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}| &= \left|\mathfrak{D}\mathfrak{I}^{ii}\mathfrak{I}^{jj} + \sum_{\nu=1}^{m} \mathfrak{U}^{(\nu)}\mathfrak{I}_{jj} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}\right| \\ &= \sum_{\nu=1}^{m} |\mathfrak{D}\mathfrak{I}^{ii}\mathfrak{I}^{jj} + \mathfrak{U}^{(\nu)}\mathfrak{I}_{jj} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}|.\end{aligned}$$

Thus, if we define the matrix $\mathfrak{H}^{(\mu\nu)}$ by

$$(4.3)\quad \mathfrak{H}^{(\mu\nu)} = \left(\frac{1-\delta_{ij}}{D}\,|\mathfrak{D}\mathfrak{I}^{ii}\mathfrak{I}^{jj} + \mathfrak{U}^{(\nu)}\mathfrak{I}_{jj} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}|\right),\ \mu, \nu = 1, 2, \ldots, m,$$

we have

$$\sum_{\nu=1}^{m} \mathfrak{H}^{(\mu\nu)} = \left(\frac{1-\delta_{ij}}{\mathfrak{D}}\,|\mathfrak{D}\mathfrak{I}^{ii} + \mathfrak{U}^{(\mu)}\mathfrak{I}_{ji}|\right).$$

Hence,

$$\mathfrak{D}^{-1}\mathfrak{U}^{(\mu)} = \sum_{\nu=1}^{m+1} \mathfrak{H}^{(\mu\nu)}.$$

An examination of (4.3) reveals that, if $\nu = \mu$, the determinantal factor of the general element of $\mathfrak{H}^{(\mu\nu)}$ has two identical columns. Thus,

$$(4.4)\qquad \mathfrak{H}^{(\mu\mu)} = \mathfrak{O}.$$

Again, interchanging the symbols μ and ν in formula (4.3) has the effect of interchanging two columns in the determinantal factor. Since this changes the sign of the determinant, we infer that

$$(4.5)\qquad \mathfrak{H}^{(\mu\nu)} = -\,\mathfrak{H}^{(\nu\mu)}.$$

This proves the lemma.

It may also be noted that

$$(4.6)\qquad \sum_{\mu=1}^{m} \mathfrak{H}^{(\mu,m+1)} = \mathfrak{I}.$$

This is obtained by summing (4.2).

In anticipation of a notational device to be introduced later, we define the matrix $\mathfrak{H}^{(m+1,\mu)}$ by

$$\mathfrak{H}^{(m+1,\mu)} = -\mathfrak{H}^{(\mu,m+1)} \tag{4.7}$$

and let $\mathfrak{H}^{(m+1,m+1)}$ be the zero matrix of order n. Relations (4.4) and (4.7) are then valid for $\mu = 1, 2, \ldots, m+1$, and relation (4.5) is valid for $\mu, \nu = 1, 2, \ldots, m+1$.

THEOREM 5. *There exist matrices* $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$, $\mu, \nu = 1, 2, \ldots, m+1$, *such that*

$$\mathfrak{G}^{(\mu)}(x, s, \lambda) = \sum_{\nu=1}^{m+1} \mathfrak{G}^{(\mu\nu)}(x, s, \lambda), \ \mu = 1, 2, \ldots, m, \tag{4.8}$$

and

$$\mathfrak{G}^{(\mu\nu)}(x, s, \lambda) = -\mathfrak{G}^{(\nu\mu)}(x, s, \lambda). \tag{4.9}$$

To prove the theorem, let the matrix $\mathfrak{U}^{(\mu)}$, appearing in Lemma 1, be specified by

$$\mathfrak{U}^{(\mu)} = \mathfrak{W}^{(\mu)}(\lambda)\mathfrak{Y}(a_\mu, \lambda), \qquad \mu = 1, 2, \ldots, m. \tag{4.10}$$

The matrix $\mathfrak{D}$ of that lemma becomes, then, the characteristic matrix $\mathfrak{D}(\lambda)$ and will be non-singular if λ is not a characteristic value. Hence,

$$\mathfrak{D}^{-1}(\lambda)\mathfrak{W}^{(\mu)}(\lambda)\mathfrak{Y}(a_\mu, \lambda) = \sum_{\nu=1}^{m+1} \mathfrak{H}^{(\mu\nu)}.$$

Let $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ be defined by

$$\mathfrak{G}^{(\mu\nu)}(x, s, \lambda) = \mathfrak{Y}(x, \lambda)\mathfrak{H}^{(\mu\nu)}\mathfrak{Y}^{-1}(s, \lambda), \qquad \mu, \nu = 1, 2, \ldots, m+1. \tag{4.11}$$

It follows at once from Lemma 1 and relation (4.7) that the relations (4.8) and (4.9) are valid.

As a particular instance of (4.9), it may be noted that

$$\mathfrak{G}^{(\mu\mu)}(x, s, \lambda) = \mathfrak{O}. \tag{4.12}$$

Further, from (4.6) we infer that

$$\sum_{\mu=1}^{m} \mathfrak{G}^{(\mu,m+1)}(x, s, \lambda) = \mathfrak{Y}(x, \lambda)\mathfrak{Y}^{-1}(s, \lambda). \tag{4.13}$$

The asymptotic representation for the solution $\mathfrak{Y}(x, \lambda)$, obtained by Langer (**1**, (6.10) and (6.11)), is

$$\mathfrak{Y}(x, \lambda) = \mathfrak{P}(x, \lambda)\mathfrak{E}(x, \lambda) \tag{4.14}$$

where,

$$\mathfrak{E}(x, \lambda) = (\delta_{ij}e^{\lambda R_j(x)}), \quad \text{with } R_j(x) = \int_{a_1}^{x} r_j(t)dt,$$

and $\mathfrak{P}(x, \lambda)$ has an asymptotic representation of the form

$$\mathfrak{P}(x, \lambda) = \mathfrak{I} + \sum_{h=1}^{k-1} \lambda^{-h}\mathfrak{P}^{(h)}(x) + \lambda^{-k}\mathfrak{B}_k(x, \lambda).$$

In the latter relation, k is any natural number and $\mathfrak{P}^{(h)}(x)$, $h = 1, 2, \ldots, k-1$, and $\mathfrak{B}_k(x, \lambda)$ are indefinitely differentiable in x, and the components of $\mathfrak{B}_k(x, \lambda)$ are analytic in λ and bounded for $|\lambda|$ large.

In view of the representation (4.14) and the definition of $\mathfrak{U}^{(\mu)}$ in (4.10), it is clear that the components of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ are exponential sums. To the end of deducing the structure of these sums, we prove the following lemma.

LEMMA 2. *The matrix* $\mathfrak{H}^{(\mu\nu)}$ *has the representation given in formulas* (4.15) *and* (4.16) *below.*

Since the components of $\mathfrak{W}^{(\mu)}(\lambda)$ are polynomials in λ, $\mathfrak{U}^{(\mu)}$ may be expressed as

$$\mathfrak{U}^{(\mu)} = (v_{ij}^{(\mu)} \exp\{\lambda R_j(a_\mu)\}),$$

where $v_{ij}^{(\mu)}$ is asymptotically a polynomial in $1/\lambda$ multiplied by some non-negative integral power of λ. $\mathfrak{H}^{(\mu,m+1)}$ is a diagonal matrix and, from (4.2), its jth diagonal component is seen to be $1/D$ multiplied by a determinant whose jth column is the jth column of $\mathfrak{U}^{(\mu)}$ and whose other columns are corresponding columns of $\mathfrak{D}$. Since $\mathfrak{D} = \mathfrak{U}^{(1)} + \mathfrak{U}^{(2)} + \ldots + \mathfrak{U}^{(m)}$, this determinant may be expanded into the sum of m^{n-1} determinants, each of which contains the jth column of $\mathfrak{U}^{(\mu)}$ as its jth column, and the αth column of one of the matrices $\mathfrak{U}^{(1)}, \mathfrak{U}^{(2)}, \ldots, \mathfrak{U}^{(m)}$ as its αth column, $\alpha \neq j$. Thus,

$$(4.15)\quad \mathfrak{H}^{(\mu,m+1)} = \left(\frac{\delta_{ij}}{D} \sum_{k_\alpha=1,\alpha\neq j}^{m} h_{\{k_\alpha|\alpha\neq j\}}^{(\mu,m+1)} \exp\left\{\lambda\left[R_j(a_\mu) + \sum_{\alpha=1,\alpha\neq j}^{n} R_\alpha(a_{k_\alpha})\right]\right\}\right),$$

where $h_{\{k_\alpha|\alpha\neq j\}}^{(\mu,m+1)}$ is asymptotically a polynomial in $1/\lambda$, multiplied by some power of λ. The subscript symbol $\{k_\alpha|\alpha \neq j\}$ is an abbreviation for the set $k_1, k_2, \ldots, k_{j-1}, k_{j+1}, \ldots, k_n$. The summation operator applies independently to each member of this set. Thus, the jth diagonal component of $\mathfrak{H}^{(\mu,m+1)}$ is exhibited as an exponential sum of m^{n-1} terms.

The matrix $\mathfrak{H}^{(\mu\nu)}$, $\mu, \nu = 1, 2, \ldots, m$, has zeros on its diagonal. From (4.3), the component in the ith row and jth column, $i \neq j$, is seen to be $1/D$ multiplied by a determinant whose ith column is the jth column of $\mathfrak{U}^{(\mu)}$, whose jth column is the jth column of $\mathfrak{U}^{(\nu)}$, and whose other columns are columns of $\mathfrak{D}$. This determinant may be expanded into the sum of m^{n-2} determinants, each of which contains, as its ith and jth columns, the jth columns of $\mathfrak{U}^{(\mu)}$ and $\mathfrak{U}^{(\nu)}$, respectively, and as its αth column, $\alpha \neq i, j$, the αth column of one of the matrices $\mathfrak{U}^{(1)}, \mathfrak{U}^{(2)}, \ldots, \mathfrak{U}^{(m)}$. Hence,

$$(4.16)\quad \mathfrak{H}^{(\mu\nu)} = \left(\frac{1-\delta_{ij}}{D} \sum_{k_\alpha=1,\alpha\neq i,j}^{m} h_{\{k_\alpha|\alpha\neq i,j\}}^{(\mu\nu)} \exp\left\{\lambda\left[R_j(a_\mu) + R_j(a_\nu) + \sum_{\alpha=1,\alpha\neq i,j}^{n} R_\alpha(a_{k_\alpha})\right]\right\}\right).$$

The notation in this relation is similar to that used in (4.15). This completes the proof of the lemma.

Recalling the definition of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ in relation (4.11) and the representation of $\mathfrak{Y}(x, \lambda)$ in (4.14), we may write

$$\mathfrak{G}^{(\mu\nu)}(x, s, \lambda) = \mathfrak{P}(x, \lambda)\mathfrak{E}(x, \lambda)\mathfrak{H}^{(\mu\nu)}\mathfrak{E}^{-1}(s, \lambda)\mathfrak{P}^{-1}(s, \lambda).$$

Anticipating the form of the product on the right, let the following two relations define their left members.

$$(4.17)\quad \phi^{(\mu,m+1)}_{[k_\alpha|\alpha\neq j]}(x, s) = R_j(x) - R_j(s) + R_j(a_\mu) + \sum_{\alpha=1,\alpha\neq j}^{n} R_\alpha(a_{k_\alpha}), \qquad \mu = 1, 2, \ldots, m,$$

$$(4.18)\quad \phi^{(\mu\nu)}_{[k_\alpha|\alpha\neq i,j]}(x, s) = R_i(x) - R_j(s) + R_j(a_\mu) + R_j(a_\nu) + \sum_{\alpha=1,\alpha\neq i,j}^{n} R_\alpha(a_{k_\alpha}), \qquad \mu, \nu = 1, 2, \ldots, m,\ \mu \neq \nu.$$

Both $\mathfrak{E}(x, \lambda)$ and its inverse are diagonal matrices, hence, multiplying each of the relations (4.15) and (4.16) on the left by $\mathfrak{E}(x, \lambda)$ and on the right by $\mathfrak{E}^{-1}(s, \lambda)$, we have

$$\mathfrak{E}(x, \lambda)\mathfrak{H}^{(\mu,m+1)}\mathfrak{E}^{-1}(s, \lambda) = \left(\frac{\delta_{ij}}{D}\sum_{k_\alpha=1,\alpha\neq j}^{m} h^{(\mu,m+1)}_{[k_\alpha|\alpha\neq j]} \exp\{\lambda\phi^{(\mu,m+1)}_{[k_\alpha|\alpha\neq j]}(x, s)\}\right)$$

and

$$\mathfrak{E}(x, \lambda)\mathfrak{H}^{(\mu\nu)}\mathfrak{E}^{-1}(s, \lambda) = \left(\frac{1-\delta_{ij}}{D}\sum_{k_\alpha=1,\alpha\neq i,j}^{m} h^{(\mu\nu)}_{[k_\alpha|\alpha\neq i,j]} \exp\{\lambda\phi^{(\mu\nu)}_{[k_\alpha|\alpha\neq i,j]}(x, s)\}\right).$$

The matrix $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ is obtained by multiplying the appropriate one of the above matrices on the left by $\mathfrak{P}(x, \lambda)$ and on the right by $\mathfrak{P}^{-1}(s, \lambda)$. In this connection, we may observe that each component of the product, $\mathfrak{A}\mathfrak{B}\mathfrak{C}$, of three matrices is a linear combination of all the components of $\mathfrak{B}$, and that each coefficient in this linear combination is the product of some component of $\mathfrak{A}$ with some component of $\mathfrak{C}$. From this, and the fact that the components of both $\mathfrak{P}(x, \lambda)$ and $\mathfrak{P}^{-1}(s, \lambda)$ are asymptotically polynomials in $1/\lambda$, it is clear that each component of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ will be an exponential sum containing, in general, all the exponential terms appearing in $\mathfrak{H}^{(\mu\nu)}$. The coefficients of these sums will, moreover, be of the same form as the coefficients in the non-zero components of $\mathfrak{H}^{(\mu\nu)}$, except that they will be functions of x and s. Hence, each component of $\mathfrak{G}^{(\mu,m+1)}(x, s, \lambda)$, $\mu = 1, 2, \ldots, m$, is of the form

$$(4.19)\qquad \lambda^{\rho}/D \sum_{j=1}^{n} \sum_{k_\alpha=1,\alpha\neq j}^{m} g^{(\mu,m+1)}_{[k_\alpha|\alpha\neq j]} \exp\{\lambda\phi^{(\mu,m+1)}_{[k_\alpha|\alpha\neq j]}(x, s)\}.$$

Similarly, each component of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$, $\mu, \nu = 1, 2, \ldots, m$, is of the form

$$(4.20)\qquad \lambda^{\rho}/D \sum_{i,j=1}^{n} \sum_{k_\alpha=1,\alpha\neq i,j}^{m} g^{(\mu\nu)}_{[k_\alpha|\alpha\neq i,j]} \exp\{\lambda\phi^{(\mu\nu)}_{[k_\alpha|\alpha\neq i,j]}(x, s)\}.$$

The non-negative integer θ is defined to be the smallest such integer for which the coefficients, $g_{[k_\alpha|\alpha\neq j]}{}^{(\mu,m+1)}$ and $g_{[k_\alpha|\alpha\neq t,j]}{}^{(\mu\nu)}$, are asymptotic polynomials in $1/\lambda$ for every admissible value of their various indices. The above results are summarized in the following theorem.

THEOREM 6. *Each component of* $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ *is an exponential sum of the form shown in* (4.19) *or* (4.20). *The coefficient of* λ *in the exponent of* e *in each term of the sum is given by* (4.17) *or* (4.18).

As a useful notational device, we define the square matrix $((\mathfrak{G}))$, whose components are matrices, by the relation

$$((\mathfrak{G})) = ((\mathfrak{G}^{(\mu\nu)}(x, s, \lambda))), \qquad \mu, \nu = 1, 2, \ldots, m + 1.$$

Because of relation (4.9) in Theorem 5, this matrix is seen to be skew-symmetric. Further, let the symmetric matrix $\mathfrak{F}$ be defined by

$$\mathfrak{F} = (\phi^{(\mu\nu)}(x, s)), \qquad \mu, \nu = 1, 2, \ldots, m + 1.$$

The components of this matrix are the functions defined in (4.17) and (4.18) for all values of μ and ν for which those definitions are valid. The definition of the remaining components is achieved by the relations

$$\phi^{(\mu\nu)}(x, s) \equiv 0, \qquad \text{if } \mu = \nu,$$

$$\phi^{(m+1,\nu)}(x, s) = \phi^{(\nu,m+1)}(x, s), \qquad \nu = 1, 2, \ldots, m + 1.$$

Thus, the element in the μth row and νth column of $\mathfrak{F}$ corresponds uniquely to the element in the μth row and νth column of $((\mathfrak{G}))$. That is to say, the exponential sum which constitutes the general component of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ is $1/D$ multiplied by a linear combination of exponential terms of the form $\exp\{\lambda\phi^{(\mu\nu)}(x, s)\}$, where the undesignated parameters in $\phi^{(\mu\nu)}(x, s)$ are allowed to range through all their admissible values. It follows that, when we are concerned with the sum of any specific block of components in $((\mathfrak{G}))$, the exponential sums contained therein will have in their exponents precisely those ϕ-functions which appear in the corresponding block of components in $\mathfrak{F}$.

The sums of certain blocks of components in $((\mathfrak{G}))$ can be concisely represented, if we define the vector $\mathfrak{b}_j$ to be an $(m + 1)$-dimensional vector with 1 in the jth place and zeros elsewhere and define the vector $\mathfrak{i}_q$ by the relation

$$\mathfrak{i}_q = \sum_{j=1}^{q} \mathfrak{b}_j.$$

Thus, recalling (4.8),

$$\mathfrak{b}_\mu((\mathfrak{G}))\mathfrak{i}_{m+1} = \mathfrak{G}^{(\mu)}(x, s, \lambda).$$

Hence,

$$\sum_{\mu=1}^{q} \mathfrak{G}^{(\mu)}(x, s, \lambda) = \mathfrak{i}_q((\mathfrak{G}))\mathfrak{i}_{m+1},$$

and it is immediately clear that this notation can be used to rewrite formula (3.1). That is,

$$(4.21)\qquad \mathfrak{G}(x, s, \lambda) = \begin{cases} \mathfrak{i}_q((\mathfrak{G}))_{m+1}, & s < x \\ -(\mathfrak{i}_m - \mathfrak{i}_q)((\mathfrak{G}))\mathfrak{i}_{m+1}, & s > x \end{cases} \; s \text{ on } (a_q, a_{q+1}).$$

THEOREM 7. *Formula* (4.21) *for Green's matrix may be reduced to*

$$(4.22)\qquad \mathfrak{G}(x, s, \lambda) = \begin{cases} \mathfrak{i}_q((\mathfrak{G}))(\mathfrak{i}_{m+1} - \mathfrak{i}_q), & s < x \\ -(\mathfrak{i}_m - \mathfrak{i}_q)((\mathfrak{G}))(\mathfrak{i}_q + \mathfrak{d}_{m+1}), & s > x \end{cases} \; s \text{ on } (a_q, a_{q+1}).$$

This result follows immediately when it is recalled that $((\mathfrak{G}))$ is skew-symmetric, and hence, that both $\mathfrak{i}_q((\mathfrak{G}))\mathfrak{i}_q$ and $(\mathfrak{i}_m - \mathfrak{i}_q)((\mathfrak{G}))(\mathfrak{i}_m - \mathfrak{i}_q)$ are zero.

The simplification of Green's matrix achieved by Theorem 7 is of basic significance. In its absence, the definition of regularity would of necessity be made in terms of formula (4.21). Such a definition would not permit the fundamental conclusion stated in Theorem 8 below.

5. Regularity of the boundary problem. In § 4 it was noted that each component of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ is $1/D$ multiplied by an exponential sum. Since D is itself an exponential sum given by (**1**, (11.3))

$$D = D(\lambda) = \sum_{\alpha} A_\alpha(\lambda) e^{\lambda \Omega_\alpha},$$

each component of $\mathfrak{G}^{(\mu\nu)}(x, s, \lambda)$ may be interpreted as the quotient of two exponential sums. A comparison of the exponents of the numerator with those of the denominator is clearly vital to a discussion of the convergence of $\mathfrak{b}_k^{(l)}(x)$ defined in (3.8).

Let the set of exponent coefficients $\{\Omega_\alpha | A_\alpha(\lambda) \not\equiv 0\}$ be represented by the symbol E_D. This set is a subset of the set E defined by

$$\left\{ \sum_{\alpha=1}^{n} R_\alpha(a_{k_\alpha}) \right\},$$

where each member of $k_1, k_2, \ldots, k_n$ is chosen independently from the integers $1, 2, \ldots, m$, (**1**, (11.3) *et seq.*). Let the members of the set E_D be plotted on a complex z-plane, and let P_D be the closed region bounded by the convex polygon of smallest area which contains all these points in its interior or on its perimeter. It may be noted for future reference that the members of the set E may be similarly plotted and that they will determine a corresponding closed minimum convex polygonal region P. The region P_D may coincide with P, but if certain members of the set $\{A_\alpha(\lambda)\}$ are identically zero, P_D will be a proper subregion of P.

The exponent coefficients, defined in (4.17) and (4.18), are functions of s for each fixed value of x and each permissible set of values of the parameters involved. If the symbol $\phi^{(\mu\nu)}(x, s)$ is used to represent any one of these functions the relation

$$(5.1)\qquad z = \phi^{(\mu\nu)}(x, s)$$

will effect a mapping of any s-interval into a complex z-plane. Since $R_j(s)$ has a constant argument and $R_j'(s) \neq 0$, the image is a straight line and the mapping is one-to-one. It may be similarly inferred that, for s fixed, the relation (5.1) will effect a one-to-one mapping of any x-interval into a straight line image. The definition of regularity will be made in terms of the location of the s-interval images relative to the region P_D defined above.

Definition. The boundary problem will be said to be regular relative to a specific value of x if, for all permissible values of the parameters $\{k_\alpha | \alpha \neq j\}$ or $\{k_\alpha | \alpha \neq i, j\}$, as the case may be:

(i) Every ϕ-function in the sum

$$\mathfrak{i}_q \mathfrak{F}(\mathfrak{i}_{m+1} - \mathfrak{i}_q)$$

maps (a_q, a_{q+1}) into P_D for every q such that $a_{q+1} < x$, and maps (a_q, x) into P_D when $a_q < x < a_{q+1}$;

(ii) Every term in the sum

$$(\mathfrak{i}_m - \mathfrak{i}_q)\mathfrak{F}(\mathfrak{i}_q + \mathfrak{d}_{m+1})$$

maps (a_q, a_{q+1}) into P_D for every q such that $a_q > x$, and maps (x, a_{q+1}) into P_D when $a_q < x < a_{q+1}$.

The boundary problem will be said to be regular relative to any subinterval of $[a_1, a_m]$, if it is regular relative to every x on that subinterval.

It will be seen, on recalling the representation of $\mathfrak{G}(x, s, \lambda)$ in (4.22), that if a problem is regular, every exponent coefficient in the exponential sum constituting the numerator of each component of $\mathfrak{G}(x, s, \lambda)$ will have values lying in P_D for all values of the variable s.

A sufficient condition for regularity will now be developed by showing that each s-interval mentioned in the definition of regularity is mapped into the region P by the mapping functions associated with it. From this it will follow that if P_D coincides with P the boundary problem is regular.

If, in the mapping relation (5.1), $\phi^{(\mu\nu)}(x, s)$ is the function defined by (4.18), it is clear that the image points $\phi^{(\mu\nu)}(a_1, a_\mu)$ and $\phi^{(\mu\nu)}(a_m, a_\mu)$ belong to the set E and are, therefore, in P. Hence, since P is convex, $\phi^{(\mu\nu)}(x, a_\mu)$ is in P for any x on $[a_1, a_m]$. Similarly, it may be inferred that $\phi^{(\mu\nu)}(x, a_\nu)$ is in P for the same x. This leads to the conclusion contained in the following lemma.

LEMMA 3. *The relation* (5.1) *with* $\mu, \nu = 1, 2, \ldots, m$, $(\mu \neq \nu)$ *maps the s-interval* $[a_\mu, a_\nu]$ *into a line in P for any fixed x on* $[a_1, a_m]$. *Moerover, if s is bounded away from the end points of its interval, z is bounded away from the vertices of P.*

If $\nu = m + 1$ in (5.1) and $\phi^{(\mu, m+1)}(x, s)$ is defined by (4.17), it is clear that the images of all pairs of values of x and s lie on the same straight line. Since the points $\phi^{(\mu, m+1)}(a_1, a_\mu)$ and $\phi^{(\mu, m+1)}(a_m, a_\mu)$ lie in P, the point $\phi^{(\mu, m+1)}(x, a_\mu)$ lies in P for any x on $[a_1, a_m]$. Noting, then, that $\phi^{(\mu, m+1)}(x, x)$ is in P, we can state the following lemma.

LEMMA 4. *The relation* (5.1) *with* $\nu = m + 1$, $\mu = 1, 2, \ldots, m$, *maps the s-interval* $[a_\mu, x]$ *into a line in* P *for any* x *on* $[a_1, a_m]$. *If* s *is bounded away from* x *and* a_μ, z *is bounded away from the vertices of* P.

Let (a_q, a_{q+1}) be any s-interval determined by a pair of consecutive boundary points. If $a_{q+1} < x$, it is readily seen, by employing the above lemmas, that relation (5.1) maps (a_q, a_{q+1}) into P provided that $\mu < q$ and $\nu > q + 1$. For, under these conditions on μ and ν, (a_q, a_{q+1}) is contained in $[a_\mu, a_\nu]$ when $\nu \neq m + 1$ and is contained in $[a_\mu, x]$ when $\nu = m + 1$. If $a_q < x < a_{q+1}$, a similar argument shows that (a_q, x) is mapped into P by (5.1) when $\mu < q$ and $\nu > q + 1$. These facts can be summarized by saying that each ϕ-function in the sum

$$\mathfrak{i}_q \mathfrak{F}(\mathfrak{i}_{m+1} - \mathfrak{i}_q)$$

maps (a_q, a_{q+1}) into P for every q such that $a_{q+1} < x$, and maps (a_q, x) into P when $a_q < x < a_{q+1}$. In a similar fashion, it can be inferred that each ϕ-function in the sum

$$(\mathfrak{i}_m - \mathfrak{i}_q)\mathfrak{F}(\mathfrak{i}_q + \mathfrak{d}_{m+1})$$

maps (a_q, a_{q+1}) into P for every q such that $a_q > x$, and maps (x, a_{q+1}) into P when $a_q < x < a_{q+1}$. Comparing these results with the definition of regularity, the following theorem can be stated.

THEOREM 8. *If* P_D *coincides with* P, *the boundary problem is regular.*

The above theorem establishes the fact that all problems in the category initially specified are regular except possibly those for which the determinant $D(\lambda)$ is degenerate in the sense that P_D is a proper subregion of P. Success in establishing this fact depended on the relations (4.9) and (4.12), by means of which the original form of Green's matrix given in (3.1) was simplified to the form exhibited in (4.22). The relations in question apply equally well in the more general complex case. Consequently, Langer's regularity conditions could, with advantage, be amplified to include a recognition of the simplifying properties of these relations. In this connection, it should be noted that Langer made specific mention of the possibility of a simplification within the formula for a single matrix $\mathfrak{G}^{(\mu)}(x, s, \lambda)$, but that the simplification suggested here occurs between the terms of a sum of such matrices. Hence, in order to take advantage of the relations, the paths of integration, corresponding to those in formula (2.4), need to be chosen so that some of them have segments in common. This will generally be possible and the attendant simplification will be sufficient to admit as regular many problems (the present one is a case in point) which would not be regular according to a literal interpretation of Langer's conditions.

6. Convergence of the expansion. The convergence discussion given in **(1)** is applicable here, but it will be replaced by one which imposes a less restrictive condition on the vector to be expanded.

The matrix $\mathfrak{G}(x, s, \lambda)$ is a sum of matrices whose components are displayed in (4.19) and (4.20). The multiplication of $\mathfrak{G}(x, s, \lambda)$ on the right by $\mathfrak{R}(s)\mathfrak{f}^{(r+1)}(s)$ will, therefore, yield a vector whose components are sums of functions of the form

$$\frac{\lambda^\theta}{D} h^{(\mu\nu)}(x, s, \lambda) \exp\{\lambda\phi^{(\mu\nu)}(x, s)\},$$

where $h^{(\mu\nu)}(x, s, \lambda)$ is asymptotically a polynomial in $1/\lambda$. Let the integer ρ be chosen as in **(1**, § 12**)** so that, for each α for which Ω_α is a vertex of the polygon bounding P_D, the function

$$\lambda^{-\rho} D e^{-\lambda\Omega_\alpha},$$

(1, (12.2)**)**, is uniformly bounded from zero for λ on the contours of the set $\{\Gamma_k\}$. Define $k^{(\mu\nu)}(x, s, \lambda)$ by

$$\frac{1}{D} h^{(\mu\nu)}(x, s, \lambda) = k^{(\mu\nu)}(x, s, \lambda)\lambda^\rho e^{\lambda\Omega_\alpha}$$

so that $k^{(\mu\nu)}(x, s, \lambda)$ is bounded and integrable in s and λ for λ on Γ_k. A typical term in the sum that comprises any component of the vector $\mathfrak{b}_k^{(l)}(x)$, as defined in (3.8), is given by

$$(6.1) \qquad \int_{\Gamma_k}\int_{a_1}^{a_m} \lambda^{l+\theta-\rho-r-1} k^{(\mu\nu)}(x, s, \lambda) \exp\{\lambda(\phi^{(\mu\nu)}(x, s) - \Omega_\alpha)\}\, ds d\lambda.$$

If x is a point at which the boundary problem is regular, $\phi^{(\mu\nu)}(x, s)$ lies in the region P_D for every s on (a_1, a_m), with the exception of the boundary points $a_2, a_3, \ldots, a_{m-1}$, and the point x at each of which the integrand is not defined. For any λ, then, the index α can be chosen so that

$$(6.2) \qquad R\{\lambda(\phi^{(\mu\nu)}(x, s) - \Omega_\alpha)\} < 0$$

for all values of s. There will, moreover, exist a sector on the λ-plane, which may be specified by

$$(6.3) \qquad \xi_\alpha < \arg \lambda < \xi_\alpha'$$

such that the inequality (6.2) is maintained for all λ therein. A finite set of such sectors will cover the whole λ-plane and will effect a subdivision of the contour Γ_k into segments. The symbol $\Gamma_{k\alpha}$ will be used to designate that segment which lies in the sector specified by (6.3). The integral (6.1) may be expressed as a sum according to the partition of Γ_k, and a further decomposition is determined by partitioning $[a_1, a_m]$ at the points $a_2, a_3, \ldots, a_{m-1}$, and x, where the integrand is discontinuous. In consequence, we may say that any component of the vector $\mathfrak{b}_k^{(l)}(x)$ consists of a sum of terms of the type

$$(6.4) \qquad \int_{\Gamma_{k\alpha}}\int_c^d \lambda^{-1}\varphi(x, s, \lambda) ds d\lambda,$$

where c and d are any two consecutive partition points of $[a_1, a_m]$ and

$$(6.5) \qquad \varphi(x, s, \lambda) = \lambda^{l+\theta-\rho-r} k^{(\mu\nu)}(x, s, \lambda) \exp\{\lambda(\phi^{(\mu\nu)}(x, s) - \Omega_\alpha)\}.$$

The non-negative integer τ, on which the expansion depends, is assumed to be at least as large as $\theta - \rho$ and sufficiently small to insure the existence of $\mathfrak{f}^{(\tau+1)}(x)$. Let the exponent of λ in (6.5) be written as $l - l_1$, where $l_1 = -\theta + \rho + \tau$. If $l \leqq l_1$, it is clear that $\varphi(x, s, \lambda)$ is bounded and integrable for λ large.

If $l < l_1$,

$$\lim_{|\lambda| \to \infty} \varphi(x, s, \lambda) = 0$$

uniformly in s on (c, d). Thus it is easily inferred that integral (6.4) converges to zero as $k \to \infty$. From this it follows that $\mathfrak{b}_k^{(l)}(x)$ converges to zero and $\mathfrak{S}_k^{(l)}(x)$ converges to $\mathfrak{f}^{(l)}(x)$ as $k \to \infty$.

If $l = l_1$ and if, for ϵ arbitrary, arg λ and s are restricted by $\xi_\alpha + \epsilon \leqq \arg \lambda \leqq \xi_\alpha' - \epsilon$ and $c + \epsilon \leqq s \leqq d - \epsilon$, respectively, then, recalling Lemmas 3 and 4, § 5,

$$\lim_{|\lambda| \to \infty} \exp\{\lambda(\phi^{(\mu\nu)}(x, s) - \Omega_\alpha)\} = 0,$$

uniformly in s. At once,

$$\lim_{|\lambda| \to \infty} \varphi(x, s, \lambda) = 0,$$

uniformly in s, for arg λ and s restricted as above. From this it follows easily (see (**2**, Lemma 1, p. 166)) that integral (6.4) converges to zero, and hence that $\mathfrak{S}_k^{(l)}(x)$ converges to $\mathfrak{f}^{(l)}(x)$ as $k \to \infty$.

Combining the two cases, then, it may be stated that the series $\mathfrak{S}^{(l)}(x)$ converges to $\mathfrak{f}^{(l)}(x)$ for $l \leqq l_1$. The convergence is readily seen to be uniform in x on any closed interval on which the boundary problem is regular.

If $l < l_1$, it is easily inferred (see (**1**, § 17)) that the series arising from the term-by-term differentiation of $\mathfrak{S}^{(l)}(x)$ converges to $\mathfrak{f}^{(l)\prime}(x)$. In particular, $\mathfrak{S}^{(0)}(x)$ converges uniformly to $\mathfrak{f}(x)$, and this series admits of term-by-term differentiation to the order l_1. The following theorem summarizes some of these results.

THEOREM 9. *Let τ be the larger of the integers* 0 *and $\theta - \rho$. If $\mathfrak{f}(x)$ is any vector with a bounded and integrable derivative of order* $(\tau + 1)$ *on a closed subinterval* $[c, d]$ *on which the boundary problem is regular, then the series expansion $\mathfrak{S}^{(0)}(x)$, associated with τ, converges uniformly to $\mathfrak{f}(x)$ on* $[c, d]$. *Moreover, if $\theta - \rho$ is negative, this series admits of term-by-term differentiation to the order of $\rho - \theta$.*

REFERENCES

1. R. E. Langer, *The boundary problem of an ordinary linear differential system in the complex domain*, Trans. Amer. Math. Soc., *46* (1939), 151–190.
2. ——— *Developments associated with a boundary problem not linear in the parameter*, Trans. Amer. Math. Soc., *25* (1923), 155–172.

3. C. E. Wilder, *Expansion problems of ordinary linear differential equations with auxiliary conditions at more than two points*, Trans. Amer. Math. Soc., *18* (1917), 415–442.

4. R. H. Cole, *Reduction of an n-th order differential equation and m-point boundary conditions to an equivalent matrix system*, Amer. J. Math. (1), *68* (1946), 179–184.

5. G. H. Haltiner, *The theory of linear differential systems based upon a new definition of the adjoint*, Duke Math. J., *15* (1948), 893–919.

6. W. M. Whyburn, *Differential systems with general boundary conditions*, Seminar Reports in Mathematics, University of California Publ. Math. (1), *2* (1944), 45–61.

7. ——— *Differential equations with general boundary conditions*, Bull. Amer. Math. Soc. (10), *48* (1942), 692–704.

8. ——— *Differential systems with boundary conditions at more than two points*, Proceedings of the Conference on Differential Equations, held at the University of Maryland, March 17–19, 1955 (University of Maryland Bookstore, College Park, Md., 1956), 1–21.

University of Western Ontario

ON STABILITY IN THE LARGE FOR SYSTEMS OF ORDINARY DIFFERENTIAL EQUATIONS

PHILIP HARTMAN

1. Autonomous systems. This note concerns the stability of systems of (real) differential equations in the large on Euclidean space E^n and on certain Riemannian manifolds M^n. The results will be refinements of those of Krasovski **(3)**, **(4)**, **(5)** and of Markus and Yamabe **(8)** and will make clear the role of the various assumptions in the type of theorems under consideration.

In this section, the main theorems are stated for autonomous systems

$$x' = f(x). \tag{1}$$

Their proofs are given in § 2, 3, 4. In § 5, 6, 7, generalizations to non-autonomous systems are made.

The following notation will be used below: Let A^* denote the transpose of the (real) matrix $A = (a_{jk})$, A^H the Hermitian part, $\frac{1}{2}(A + A^*)$, of A. For any two matrices A and B, let $A < B$ mean that $A^H < B^H$, that is, that $B^H - A^H$ is positive definite. Finally, let I be the unit matrix. For points x, y of Euclidean space, $x \cdot y$ denotes the scalar product and $|x| = (x \cdot x)^{\frac{1}{2}} \geqslant 0$. It will generally be assumed that:

(A) $M = M^n$ is a complete Riemannian manifold with a positive, definite, metric tensor $g_{ik}(x)$ of class C^1, and $f(x)$ is a contravariant vector field of class C^1 on M. (The covariant derivative of f is the tensor with components

$$f^k{}_{,m} = \partial f^k/\partial x^m + g^{ik}[jm, i]f^j,$$

where

$$[jm, i] = \tfrac{1}{2}(\partial g_{ji}/\partial x^m + \partial g_{mi}/\partial x^j - \partial g_{jm}/\partial x^i).)$$

The distance between two points x, y of M, considered as a metric space, will be denoted by $d(x, y)$. By $d(x)$ will be meant the distance $d(x, x^0)$ from x to a fixed point x^0 of M.

LEMMA 1. *Assume* (A). *Suppose that the tensor $e_{ij} = g_{ik}f^k{}_{,j}$ satisfies*

$$(e_{ij}) < 0. \tag{2}$$

Received June 28, 1960. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under contract No. AF 18(603)-41. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Then every solution $x = x(t)$ of (1) *exists for large $t > 0$; furthermore, if $x = x_1(t)$, $x_2(t)$ are two distinct solutions of* (1) *for $t > 0$, then*

(3) $$d(x_1(t), x_2(t)) \text{ is decreasing}$$

for $t > 0$. In particular, if there exists a stationary point $x = x_0$,

(4) $$f(x_0) = 0,$$

then every solution $x = x(t) \not\equiv x_0$ satisfies

(5) $$d(x_0, x(t)) \downarrow 0, t \to \infty$$

(where "$\downarrow$" signifies "decreasing").

It can be remarked that if the condition (2) is relaxed to

(2′) $$(e_{ij}) < 0,$$

then the assertion concerning the existence of $x(t)$ for large t remains valid, but (3) must be replaced by

(3′) $$d(x_1(t), x_2(t)) \text{ is non-increasing}$$

and, of course, (4) then does not imply (5) or even $d(x(t), x_0) \to 0$ as $t \to \infty$. Assertion (3′) implies, however, that there is a constant C, depending only on $f(x)$ with the property that if $x = x(t)$ is any solution of (1) for $t > T$, then

(6) $$d(x(t)) < d(x(T)) + C(t - T) \text{ for } t > T.$$

In order to see this, let $x = x_1(t)$ be the solution of (1) satisfying $x_1(0) = x^0$, where x^0 is the reference point of M in the definition of $d(x) = d(x, x^0)$. Let $s > 0$, and consider the solution $x = x_1(t + s)$ of (1). Then, by (3′),

$$d(x_1(t + s), x_1(t)) < d(x_1(s), x_1(0)) \text{ for } t > 0.$$

This clearly implies the existence of a constant $C > 0$ such that $d(x_1(t)) < Ct$ for $t > 0$. The inequality (6) follows from this fact and (3′), where $x(t) = x_2(t)$.

Lemma 1 is similar to results of Lewis **(7)** and Opial **(9)**. These authors deal with the case where M is replaced by a compact set. One new feature of Lemma 1 is the important remark that (2′) implies that all solutions exist for large t. The end of the proof of Lemma 1 is similar to an argument of LaSalle **(6)**.

In the last part of Lemma 1, (2) need not be required at $x = x_0$.

A consequence of (5) is that $f(x) \not\equiv 0$ for $x \not\equiv x_0$; that is, under the condition (2), there is at most one stationary point. It is of interest to note that a strengthened form of condition (2) implies the existence of a (unique) stationary point. This is the assertion of the following theorem.

(I) *Assume* (A). *Let $\lambda(r)$ be a positive, non-increasing function of r for $r > 0$ such that*

(7) $$\int^\infty \lambda(r) dr = \infty.$$

Let the tensor $e_{jk} = g_{jm}f^m{}_{,k}$ *satisfy*

(8) $$(e_{jk}(x)) < -\lambda(d(x))(g_{jk}(x)).$$

Then there exists a unique point $x = x_0$ *of* M *satisfying* $f(x_0) = 0$. (*Hence, by Lemma* 1, *all solutions* $x = x(t) \not\equiv x_0$ *of* (1) *satisfy* (5).)

Markus and Yamabe **(8)**[1] prove a result concerning solutions of (1) in which it is assumed that f satisfies (8), but (7) is replaced by the stronger condition

$$\int^{\infty}\left[\exp(-\epsilon)\int_0^t \lambda(u)du\right]dt < \infty \qquad \text{for all } \epsilon > 0.$$

Although their assumption is stronger than (7), their conclusion is apparently weaker than (5), since they did not notice (3) or that (1) has a *unique* stationary point. For weaker versions of (I) in the case that M^n is Euclidean space E^n (or the vector space R^n with a constant positive definite metric tensor $G = (g_{jk})$), see **(3)**, **(4)**, **(5)**.

If the proof of (I) is combined with that of Lemma 1, there results the estimate

$$d(x(t), x_0) < d(x(S), x_0)e^{-\lambda(c)(t-S)}$$

for $t > S$ if $x(t)$ is defined for $t = S$. In this inequality, $c = d(x_0) + d(x(S), x_0)$.

It turns out that most of the assertions of (I) remain valid if (8) is relaxed to

(9) $$e_{jk}f^jf^k < -\lambda(d(x))g_{jk}f^jf^k.$$

(Ia) *Assume all conditions of* (I) *except that* (8) *is replaced by* (9). *Then:*

(i) *any solution* $x = x(t)$ *of* (1) *defined at* $t = 0$ *exists for* $t > 0$;

(ii) *the limit* $x(\infty) = \lim x(t)$, $t \to \infty$, *exists and is a stationary point*, $f(x(\infty)) = 0$;

(iii) *if* $x(t) \not\equiv x(\infty)$ *and*

$$v(t) = (g_{jk}f^jf^k)^{\frac{1}{2}} \text{ at } x = x(t),$$

then $v(t) \downarrow 0$, $t \to \infty$;

(iv) *the set of stationary points* $x = x_0$ *of* $f(x)$ *is connected*; *hence*,

(v) *if the stationary points* $x = x_0$ *of* $f(x)$ *are isolated* (*for example, if* $\det(e_{jk}(x_0)) \neq 0$ *whenever* $f(x_0) = 0$), *then* $f(x)$ *has a unique stationary point* $x = x_0$ (*so that* $x(\infty) = x_0$ *is independent of the particular solution* $x(t)$).

The proof of (Ia) gives the following improvements of (i)–(iii): a solution $x = x(t)$, $t > 0$, of (1) has the *a priori* bound

(10) $$d(x(t)) < c, \text{ where } c = L_1(L(d(x(0))) + v(0))$$

and $w = L_1(r)$ is the inverse function of

(11) $$L(w) = \int_0^w \lambda(r)dr;$$

also, $d(x(t)) < c$ implies

[1]Added in proof. See also Osaka Math. J., *12* (1960), 305–317.

(12) $$0 < v(t) < v(0)e^{-\lambda(c)t} \text{ for } t > 0$$

and

(12′) $$d(x(t), x(\infty)) < (v(0)/\lambda(c))e^{-\lambda(c)t} \text{ for } t > 0.$$

If condition (7) does not hold, but the initial point $x = x(0)$ of a particular solution $x = x(t)$ is such that the definition of c in (10) is meaningful, then assertions (i)–(iii) are valid for this $x = x(t)$.

The following example shows the need for the additional hypothesis in part (v) of (Ia): Let $n = 2$ and $M = E^2$ be the Euclidean plane with coordinates $x = (x^1, x^2)$. The system of differential equations $x' = -(x^1, 0)$ satisfies the analogue of (9) with $\lambda(r) \equiv 1$. The stationary points of this system form the line $x^1 = 0$. The general solution is

$$x = (x_0^1e^{-t}, x_0^2) \to (0, x_0^2)$$

and

$$v(t) = |x_0^1|e^{-t} \downarrow 0, \text{ as } t \to \infty.$$

Lemma 1 implies the following statement for the case that $M = M^n$ is the Euclidean space E^n.

LEMMA 1′. *Let $f(x)$ be of class C^1 on E^n and let $J(x) = (\partial f/\partial x)$ be the Jacobian matrix of f. Let $J(x) < 0$ for all $x \neq x_0$, where x_0 is a stationary point, $f(x_0) = 0$. Then every solution $x = x(t) \not\equiv x_0$ of* (1) *satisfies $|x(t) - x_0| \downarrow 0$, as $t \to \infty$.*

The following is a corollary of (I) when $M = E^n$ is Euclidean space.

(I′) *Let a map $T: E^n \to E^n$ be given by $y = f(x)$, where $f(x)$ is of class C^1 on E^n, and let $J(x) = (\partial f/\partial x)$. If $J(x) \leq -\lambda(|x|)I$, where $\lambda = \lambda(r)$ is as in* (I), *then T is one-to-one and onto. (Hence all solutions of* (1) *satisfy $|x(t)-x_0| \downarrow 0$ as $t \to \infty$, where $x = x_0$ is the unique point satisfying $f(x_0) = 0$.)*

It is clear that $J \leq -\lambda(|x|)I$ does not imply that T is onto (even in the case $n = 1$) if (7) fails to hold.

2. Proof of Lemma 1. Let $x = x(t; x_1)$ be the unique solution of (1) satisfying the initial condition $x(0; x_1) = x_1$. Let $x_1(t) = x(t; x_1)$ and $x_2(t) = x(t; x_2)$, where x_1, x_2 are distinct arbitrary points of M. Suppose that $x_1(t)$ exists on a closed interval $[0, T]$ where $T > 0$. Let $x = z(u)$, where $0 \leq u \leq d = d(x_1, x_2)$, be a geodesic of minimal length satisfying $z(0) = x_1$ and $z(d) = x_2$. Finally, let $x = x(t, u) = x(t; z(u))$ be the solution of (1) determined by $x(0, u) = z(u)$.

Let ϵ have the property that if $0 \leq u < \epsilon \leq d$, then $x(t, u)$ is defined for $0 \leq t \leq T$. In any case, $x(t, \epsilon)$ exists on some interval $[0, S]$. Let $L(t)$ denote the length of the curve $x = x(t, u)$, where $0 \leq u \leq \epsilon$, for a fixed t, $0 \leq t \leq S$. Then

(13) $$L(t) = \int_0^\epsilon (g_{jk}(x)y^jy^k)^{\frac{1}{2}}du,$$

where $x = x(t, u)$ and $y = \partial x(t, u)/\partial u$.

By (1), y is a solution of

(14) $$y' = J(x)y, \text{ where } J(x) = (\partial f/\partial x),$$

$x = x(t, u)$, and u is fixed. Note that $y(0, u) = \partial x(0, u)/\partial u = \partial z/\partial u \neq 0$; hence $y(t, u) \neq 0$. By (13) and (14), $L' = dL/dt$ is the integral of the product of $\frac{1}{2}(g_{jk}(x)y^jy^k)^{-\frac{1}{2}}$ and of $(g_{jk}y^jy^k)'$. This last factor is

$$(\partial g_{jk}/\partial x^m)f^my^jy^k + 2g_{jk}y^j(\partial f^k/\partial x^m)y^m.$$

If $[ij, k]$ denotes the Riemann–Christoffel symbol of the first kind, then $\partial g_{jk}/\partial x^m = [jm, k] + [km, j]$ and $\partial f^k/\partial x^m = f^k{}_{,m} - g^{ik}[jm, i]f^j$. Hence, the expression in the last formula line is $2g_{jk}f^k{}_{,m}y^ky^m$; that is,

$$(g_{jk}y^jy^k)' = 2e_{jk}y^jy^k.$$

Thus, $L'(t) < 0$ for $0 < t < S$, so that $L(t) < L(0) = d(x_1, z(\epsilon))$.

Since $d(x_1(t), x(t, \epsilon)) \leqq L(t)$,

(15) $$d(x_1(t), x(t, \epsilon)) < d(x_1, z(\epsilon))$$

for $0 < t \leqq S$. Clearly, (15) implies that the solution $x = x(t, \epsilon)$ of (1) exists for $0 \leqq t \leqq T$. Hence $x(t, u)$ exists for $0 \leqq t \leqq T$ for each fixed u, $0 \leqq u \leqq d$. In particular, $x = x_2(t) = x(t, d)$ exists for $0 \leqq t \leqq T$.

If the point x_2 in the last argument is chosen to be $x_2 = x_1(T)$, so that $x_2(t) \equiv x(t + T; x_1)$, it follows that $x_1(t)$ exists for $0 \leqq t \leqq 2T$. Repetitions of this argument show that $x_1(t)$ is defined for all $t \geqq 0$. Since x_1 is an arbitrary point of M, the first assertion of Lemma 1 follows. The second assertion (3) follows from the case $\epsilon = d$ of (15).

As to the third assertion, let $x = x(t) \not\equiv x_0$ be defined for $t \geqq 0$. Then, by (3), $d_0 = \lim d(x_0, x(t))$ exists as $t \to \infty$. Suppose, if possible, that (5) does not hold, so that $d_0 > 0$. Then there are t-values $t_1 < t_2 < \ldots.$ such that $t_m \to \infty$ and $x_1 = \lim x(t_m)$ exists, as $m \to \infty$. Clearly, $d(x_1, x_0) = d_0 > 0$. Let $x = x_m(t) = x(t - t_m; x_1)$ be the solution of (1) determined by the initial condition $x_m(t_m) = x_1$. Then $d(x_m(t), x_0) < d_0$ for $t > t_m$. The continuous dependence of solutions on initial conditions implies, therefore, that $d(x(t_m + 1), x_0) < d_0$ for large m. But this contradicts $d_0 < d(x(t), x_0) \to d_0$, $t \to \infty$. Thus Lemma 1 is proved.

3. Proof of (I)–(Ia). Let $x = x(t)$ be a solution defined at $t = 0$ and let $y = x'(t)$. Then $y = y(t)$ satisfies the linear equation (14), where $x = x(t)$. Consider the speed

(16) $$v(t) = (g_{ik}y^iy^k)^{\frac{1}{2}}, \text{ where } y = x' = f.$$

It follows that $dv^2/dt = 2e_{km}(x)y^ky^m$; see the calculation of $L'(t)$ in § 2. Thus (8) or (9) implies $dv^2/dt \leqq -2\lambda(d(x))v^2$ or, since $v > 0$,

(17) $$v' \leqq -\lambda(d(t))v, \text{ where } d(t) = d(x(t)).$$

Define a function $w = w(t)$ by

(18) $$w(t) = d(0) + \int_0^t v(s)ds.$$

By the definition of distance on M and the triangular inequality,

$$d(t) \leqq w(t), \tag{19}$$

and so, by the monotony of λ, $\lambda(d(t)) \geqq \lambda(w(t))$. Since $w' = v \geqq 0$ and $w'' = v'$, (17) implies that

$$w''(t) \leqq -\lambda(w(t))w'(t).$$

Hence

$$w'(t) \leqq w'(0) - \int_{w(0)}^{w(t)} \lambda(w)dw.$$

In view of $w' = v$ and the definition of $L(w)$ in (11), this can be written as

$$v(t) \leqq v(0) + L(d(0)) - L(w(t)).$$

Since $v(t) \geqq 0$, (19) implies that

$$L(d(t)) \leqq L(w(t)) \leqq L(d(0)) + v(0).$$

This shows that $x = x(t)$ is defined for all t and satisfies (10).

By (10) and (17), $v' \leqq -\lambda(c)v \leqq 0$ for $t \geqq 0$. Hence (12) holds and either $v(t) \equiv 0$ or $v(t) \downarrow 0$ as $t \to \infty$. Thus if $x = x_0$ is any cluster point of $x = x(t)$, $t \to \infty$, then (16) shows that $f(x_0) = 0$. In view of Lemma 1, this completes the proof of (I).

Also assertions (i), (iii) of (Ia) have been proved. The definition (16) of $v(t)$ and the inequality (12) show that the length of the curve $x = x(t)$, $0 \leqq t < \infty$, is finite,

$$\int_0^\infty (g_{jk}(x(t))x^{j'}(t)x^{k'}(t))^{\frac{1}{2}}dt = \int_0^\infty v(t)dt < \infty.$$

This implies (ii) in (Ia).

Since (v) follows from (iv), it only remains to prove (iv). The verification of (iv) to follow can be modified to show that the set of stationary points of (1) is a retract of M.

In order to prove (iv), let Q be the set of stationary points of (1). Consider a map $P: M \to Q$ defined as follows: if $x = x(t)$ is an arbitrary solution of (1) for $t \geqq 0$, put $Px(0) = x(\infty)$. It is clear that the range of P is the set Q. Since M is connected, it will follow that Q is connected if it is verified that P is continuous.

To this end, let x_1 be any point of M and M_δ the sphere $d(x_1, x) \leqq \delta$. The proof of the existence of $x(\infty)$ above shows that if $\epsilon > 0$, then there exists a $T = T(\epsilon) > 0$ independent of δ, $0 < \delta \leqq 1$, with the property that if $x(0)$ is in M_δ, then $d(x(T), x(\infty)) < \epsilon$; cf. (12'). With $T = T(\epsilon)$ fixed, choose a positive $\delta = \delta(\epsilon) \leqq 1$ so small that $d(x_1(T), x(T)) < \epsilon$ if $x = x_1(t)$, $x(t)$ are solutions of (1) determined by $x_1(0) = x_1$ and any point $x(0)$ of M_δ, respectively. Thus $x(0)$ in M_δ implies that $d(x_1(\infty), x(\infty)) < 3\epsilon$. This proves the continuity of P and completes the proof of (iv) and of (Ia).

4. On flat metrics. The proofs of Lemma 1 and (I) are particularly simple if M^n is a real n-dimensional vector space with a metric $G = \|g_{jk}\|$, where G is a constant, symmetric, positive definite matrix. If $J[s] = J(x_2s+x_1(1-s))$, then

$$f(x_2) - f(x_1) = \left(\int_0^1 J[s]ds\right)(x_2 - x_1).$$

Hence, for any constant matrix G,

$$(x_2 - x_1)\cdot G(f(x_2) - f(x_1)) = \int_0^1 (x_2 - x_1)\cdot GJ[s](x_2 - x_1)ds.$$

For example, if $GJ < 0$ and $x_1 \neq x_2$, then the integral is negative so that the map $T: M^n \to E^n$ given by $y = f(x)$ is one-to-one.

If $f(x_0) = 0$, then (1) can be written as $(x - x_0)' = f(x) - f(x_0)$. Hence $GJ < 0$ implies

$$(x - x_0)'\cdot G(x - x_0) = \int_0^1 (x - x_0)\cdot GJ[s](x - x_0)ds < 0, \tag{20}$$

for $x \neq x_0$. A simple direct proof of Lemma 1′ follows at once from this.

The equation in (20) does not seem to have been exploited in the study of stability; cf. the comparatively complicated proof in **(1)**, pp. 31-32, of the result of Krasovski which results if $GJ < 0$ is replaced by the stronger assumption $GJ(x) < -\epsilon I < 0$ and (5) by the weaker assertion $x(t) \to x_0$, $t \to 0$.

Another application of (20) will be given for non-autonomous systems in (II′) in the next section.

5. Non-autonomous systems. The results above can be generalized somewhat to systems in which t occurs explicitly,

$$x' = f(t, x). \tag{21}$$

Below it will be assumed that

(B) M, $g_{ik}(x)$, $d(x, y)$, $d(x)$ are as in (A). $f(t, x)$ is a C^1 contravariant vector field on M for every fixed $t > 0$; also f and its derivatives along M are continuous in (t, x).

The techniques of § 2 above (cf. **(7)**, **(9)**, **(10)**) imply the following analogue of Lemma 1.

LEMMA 2. *Assume* (B). *Let x_0 be a point of M satisfying*

$$f(t, x_0) = 0 \text{ for } t > 0. \tag{22}$$

Let the tensor $e_{jk}(t, x) = g_{jm}(x)f^m{}_{,k}(t, x)$ satisfy

$$(e_{jk}(t, x)) \leq 0 \ [\text{or } < 0]. \tag{23}$$

Then all solutions $x = x(t)$ of (21) *exist for large t, and $d(x_0, x(t))$ is non-increasing* [*or decreasing*]. *If, in addition, for every $c > 0$, there is a non-negative function $\mu(t) = \mu_c(t)$ defined for $t > 0$ and satisfying*

(24) $$\int^{\infty} \mu(t)dt = \infty$$

and

(25) $$(e_{jk}(t, x)) \leqq -\mu_c(t)(g_{jk}(x)) \text{ for } t \geqq 0 \text{ and } d(x) \leqq c,$$

then every solution $x = x(t)$ *of* (1) *satisfies*

(26) $$d(x_0, x(t)) \to 0 \text{ as } t \to \infty.$$

For the case that $\mu_c(t) > 0$ is independent of c and t, and $g_{jk}(x)$ is independent of x, see **(20)**, and Winter **(10)**; also Krasovski, see **(1**, p. 31**)**.

The obvious way to generalize (I) from (1) to (21) is to require an analogue of (7), (8) for a monotone $\lambda(r)$ and to assume that the length of f, $(g_{jk}(x) f^j(t, x)f^k(t, x))^{\frac{1}{2}}$, is a non-increasing function of t for every fixed x on M. But if, for example, $e_{jk}(t, x)$ satisfies

(27) $$(e_{jk}(t, x)) \leqq -\lambda(d(x))(g_{jk}(x))$$

at $t = 0$, where $\lambda = \lambda(r)$ is as in (I), then it follows that there is a unique $x = x_0$ satisfying $f(0, x_0) = 0$. Thus, when the length of f is a non-increasing function of t, one has trivially that $f(t, x_0) = 0$ for $t \geqq 0$.

A different generalization of (I) is given by

(II) *Assume* (B) *and that* $f_t = \partial f/\partial t$ *exists and is continuous in* (t, x). *Let* $\lambda(r)$ *be as in* (I) *and put*

(28) $$L(w) = \int_0^w \lambda(r)dr.$$

Let $a(t)$ *be a non-negative, continuous function integrable over* $0 \leqq t < \infty$,

(29) $$A = \int_0^{\infty} a(t)dt < \infty.$$

Let $N(w)$ *be a non-decreasing function of* w *for* $w \geqq 0$ *satisfying*

(30) $$L(w) - AN(w) \to \infty \text{ as } w \to \infty.$$

Assume that $e_{jk}(t, x)$ *satisfies* (9) *and that the length of* $f_t(t, x)$ *satisfies*

(31) $$0 \leqq [g_{jk}(x)f^j{}_t(t, x)f^k{}_t(t, x)]^{\frac{1}{2}} \leqq a(t)N(d(x))$$

for $t \geqq 0$ *and* x *in* M. *Then:*

(i) *the limit* $f(x) = \lim f(t, x)$, $t \to \infty$, *exists uniforml yon compact subsets of* M;

(ii) *every solution* $x = x(t)$ *of* (21) *exists for large* t *and tends to a limit point* $x(\infty)$ *which satisfies* $f(x(\infty)) = 0$;

(iii) *the function*

$$v(t) = (g_{jk}x^{j\prime}x^{k\prime})^{\frac{1}{2}}$$

tends to 0 *as* $t \to \infty$;

(iv) *if, in addition, there is a positive function* $\nu = \nu(c)$ *for* $c > 0$ *such that*

$$(e_{jk}(t, x)) < -\nu(c)(g_{jk}(x)) \text{ for } t > 0,\ d(x) < c,$$

then the limit function $f(x)$ *has a unique zero* $x = x_0$ *(so that* $x(\infty) = x_0$ *does not depend on the solution* $x = x(t)$*).*

The proof will furnish *a priori* bounds for $d(x(t))$ and *a priori* estimates for the $o(1)$-functions $d(x(t), x(\infty))$ and $v(t)$ depending only on the initial conditions for $x = x(t)$.

One of the main difficulties in the proof of (iv) in (II) is the fact that the limit function $f(x)$ need not be of class C^1 or even Lipschitz continuous, so that, *a priori*, it is not clear that the solutions of $x' = f(x)$ are locally unique. Local uniqueness will be proved by the use of a theorem of van Kampen **(2)**. In any case, the assertion (iv) in (II) cannot be obtained from (Ia).

If all assumptions of (II) hold except (30) and if $N(w) <$ const. $L(w)$, then (II) becomes applicable when $0 < t < \infty$ is replaced by $T < t < \infty$ for a sufficiently large T (since A is then replaced by an arbitrarily small constant).

Under the assumptions of (II), it follows that $f(t, x)$ is a bounded function of t for fixed x. This suggests the following:

(II′) *Assume* (B) *and that* $M = E^n$. *Let* G *be a positive definite, constant matrix and* $\lambda = \lambda(r)$ *a positive, non-increasing function of* $r(> 0)$. *Suppose that* α, β *are positive constants satisfying* $\alpha^2 I < G < \beta^2 I$, *that*

$$GJ(t, x) < -\lambda(|x|)I, \tag{32}$$

that $f(t, 0)$ *is a bounded function of* $t > 0$, *and that*

$$\infty > (\alpha/\beta^3) \limsup_{r\to\infty} \lambda(r)r > \underset{0<t<\infty}{\text{l.u.b.}} |f(t, 0)|. \tag{33}$$

Then every solution $x = x(t)$ *of* (21) *exists for large* t *and is bounded as* $t \to \infty$.

It will also be clear from the proof that if, in addition, either $f(t, 0) \to 0$ as $t \to \infty$ or

$$\int_0^\infty |f(t, 0)|dt < \infty, \tag{34}$$

then $|x(t)| \to 0$ as $t \to \infty$. Furthermore, if conditions (32) and (33) are replaced by the assumptions $GJ(t, x) < 0$ and (34), then the conclusions of (II′) remain valid and $\lim x(t) \cdot Gx(t)$ exists as $t \to \infty$; cf. (47) below.

6. Proof of (II). The first part of the proof of (II) is similar to that of (I). Let $x = x(t)$ be a solution of (21) on some interval $(0 <)\ S < t < T$. Define $v = v(t)$ by (16). Then

$$(v^2)' = 2e_{km}(t, x)y^k y^m + 2g_{jk}f^j f^k{}_t.$$

By Schwarz's inequality,

$$|g_{jk}f^jf^k{}_{,i}| < |g_{jk}f^jf^k|^{\frac{1}{2}}\,|g_{jk}f^j{}_{,i}f^k{}_{,i}|^{\frac{1}{2}}.$$

Thus, by (16), (9), and (31),

$$v' < -\lambda(d(t))v + a(t)N(d(t)), \text{ where } d(t) = d(x(t)). \tag{35}$$

Define $w = w(t)$ by (18), so that (19) holds and

$$w'' < -\lambda(w)w' + a(t)N(w).$$

A quadrature over $[S, t]$ gives

$$w'(t) < C - L(w(t)) + AN(w(t)),$$

where $C = w'(S) + L(w(S))$ and the justification for the last term is the fact that $N(w)$, $w(t)$ are non-decreasing in w, t, respectively.

Since $w' = v > 0$, it is clear from (30) that there does not exist any $T_0(< \infty)$ such that $w(t) \to \infty$ as $t \to T_0 - 0$. Hence $x(t)$ exists for all $t > S$ and is bounded; in fact, $d(x(t)) < c$ for $t > S$ if $L(c) - AN(c) > C$.

Let $d(t) < c$ for $t > S$, then (35) gives

$$v' < -\lambda(c)v + N(c)a(t).$$

Hence, for $t > S$,

$$0 < v(t) < v(S)e^{-\lambda(c)(t-S)} + N(c)\int_S^t e^{\lambda(c)(s-t)}a(s)\,ds, \tag{36}$$

so that (29) implies $v(t) \to 0$ as $t \to \infty$. Thus, the definition of v shows that

$$f(t, x(t)) \to 0 \text{ as } t \to \infty. \tag{37}$$

Integrating (36) for $S < t < T$ gives

$$\int_S^T v\,dt < (v(S)/\lambda(c))(1 - e^{-\lambda(c)(T-S)}) + N(c)\int_S^T e^{-\lambda(c)t}\int_S^t e^{\lambda(c)s}a(s)\,ds\,dt.$$

An integration by parts shows that the last (iterated) integral is $1/\lambda(c)$ times

$$-e^{-\lambda(c)T}\int_S^T e^{-\lambda(c)t}a(t)\,dt + \int_S^T a(t)\,dt;$$

hence

$$\lambda(c)\int_S^\infty v\,dt < v(S) + N(c)\int_S^\infty a(t)\,dt < \infty.$$

Consequently, $x(\infty) = \lim x(t)$, $t \to \infty$, exists and satisfies

$$\lambda(c)d(x(t), x(\infty)) < v(t) + N(c)\int_t^\infty a(s)\,ds.$$

The assertion (i) of (II) concerning the existence and uniformity of the limit $f(x) = \lim f(t, x)$, $t \to \infty$, is clear from (29) and (31). Furthermore, (37) implies that $f(x(\infty)) = 0$. Thus (i)–(iii) are proved.

In order to prove (iv), it is sufficient to verify the following:

(*) *Assume the conditions of* (II), *including those of* (iv) *concerning* $\nu = \nu(c)$. *Let p denote a point of M. Then solutions of*

$$p' = f(p) \tag{38}$$

are uniquely determined by initial conditions; all solutions exist for large $t > 0$; and $d(p_1(t), p_2(t))$ is a decreasing function of t if $p = p_1(t), p_2(t)$ are distinct solutions on a common t-interval.

To this end, let $x = x_1(t)$ and $x = x_2(t)$ be two distinct solutions of (21) for $t \geqq S$. Let $z = z(u)$, where $0 \leqq u \leqq d$, be a geodesic of minimal length joining $x = x_1(S), x_2(S)$ and let $x = x(t, u)$ be the solution of (21) determined by $x(S, u) = z(u)$. As in § 2, define $L(t) = L_\epsilon(t)$ by (13) for $t \geqq S, 0 \leqq \epsilon \leqq d$. Then $(e_{jk}(t, x)) \leqq 0$ implies that $L_\epsilon(t)$ is a non-increasing function of t. Since $x = x_1(t)$ is bounded as $t \to \infty$, it follows that there exists a constant $c > 0$ such that $d(x(t, u)) \leqq c$ for $t \geqq S$, $0 \leqq u \leqq d$. Hence, by (27),

$$dL_\epsilon(t)/dt \leqq -\nu(c)L_\epsilon(t) \text{ for } t \geqq S;$$

cf. the derivation of $L'(t) < 0$ in § 2. Since $d(x_1(t), x_2(t)) \leqq L_d(t)$,

$$d(x_1(t), x_2(t)) \leqq d(x_1(S), x_2(S))e^{-\nu(c)(t-S)} \tag{39}$$

for $t \geqq S$.

Let M_1 be a bounded (open) set of M. Consider the family of solutions $x = x(t; t_0, x_1)$ of (1) determined by the initial condition $x(t_0; t_0, x_1) = x_1$, where $t_0 \geqq 0$ and x_1 is a point of M_1. Then the derivation of (39) shows that there is a constant $c = c(M_1)$ such that $d(x(t; t_0, t_1)) \leqq c$ for $t \geqq t_0 \geqq 0$. Hence (39) holds for $t_0 \leqq S \leqq t < \infty$ if $x_1(t) = x(t; t_0, x_1), x_2(t) = x(t; t_0, x_2)$, and x_1, x_2 are points of M_1.

Let $y = \partial x(t; t_0, x_1)/\partial u$, where $u \neq t_0$ is one of the parameters determining the solution $x = x(t; t_0, x_1)$. Then the length of y, $(g_{jk}y^j y^k)^{\frac{1}{2}}$, is a decreasing function of t ($\geqq t_0$); cf. the derivation of (39). In particular, $y(t; t_0, x_1)$ is uniformly bounded for $t \geqq t_0$ and x_1 in M. Consequently, $x(t; t_0, x_1)$ is uniformly bounded and uniformly Lipschitz continuous with respect to t and x_1 for $t \geqq t_0 \geqq 0$ and x_1 in M_1.

It follows that there is a sequence of t-values $t_1 < t_2 < \ldots$ such that $t_n \to \infty$ and

$$p(t; x_1) = \lim_{n\to\infty} x(t + t_n; t_n; x_1) \tag{40}$$

exists uniformly for x_1 in M_1 and bounded $t \geqq 0$. Furthermore, (40) is uniformly Lipschitz continuous with respect to x_1 in M_1 for $t \geqq 0$. Note that $p_n = x(t + t_n; t_n, x_1)$ is a solution of the initial value problem

$$p_n' = f(t + t_n, p_n), \quad p_n(0) = x_1.$$

Hence (40) is a solution of

$$p' = f(p) \text{ and } p(0) = x_1. \tag{41}$$

Also, an obvious limit process in (39) shows that

(42) $d(p(t; x_1), p(t; x_2))$ is decreasing in t if $x_1 \neq x_2$,

$t > 0$ and x_1, x_2 in M_1.

Through any point of M, there passes at most one path of the family $x = p(t; x_1)$; that is,

$$p(t + s; x_1) = p(t; p(s; x_1)). \tag{43}$$

In order to prove this, let $y = y(t; t_0, x_1)$ be defined by

$$y = dx(t + t_0; t_0, x_1)/dt_0. \tag{44}$$

Then $y = x' + \partial x/\partial t_0$, and so

$$y' = x'' + J(\partial x/\partial t_0) = Jx' + f_t + J(\partial x/\partial t_0).$$

Consequently,

$$y' = Jy + f_t, \quad y(0) = 0, \tag{45}$$

where the argument of J and f_t is $(t + t_0, x(t + t_0; t_0, x_1))$. The condition $y(0) = 0$ in (45) is clear from $x(t_0; t_0, x_1) = x_1$. If $Y = (g_{jk}y^j y^k)^{\frac{1}{2}}$ is the length of y, then

$$(Y^2)' = 2e_{jk}y^j y^k + 2g_{jk}y^j f^k{}_t.$$

The derivation of (35) shows that

$$Y' < -\lambda(c)Y + N(c)a(t + t_0) < N(c)a(t + t_0).$$

Since $Y(0) = 0$,

$$Y(t) < N(c)\int_{t_0}^{t+t_0} a(s)ds \to 0 \text{ as } t_0 \to \infty.$$

As Y is the length of y, (44) implies that

$$x(t + s + t_0; s + t_0, x_1) - x(t + t_0; t_0, x_1) \to 0 \text{ as } t_0 \to \infty \tag{46}$$

uniformly for x_1 in M_1 and bounded $s, t > 0$. The relation

$$x(t + s + t_0; t_0, x_1) = x(t + s + t_0; s + t_0, x(s + t_0; t_0, x_1)),$$

the uniform Lipschitz continuity of $x(t; t_0, x_1)$ with respect to x_1, and (40) give

$$x(t + s + t_n; t_n, x_1) = x(t + t_n; t_n, p(s; x_1)) + o(1),$$

where $o(1) \to 0$ as $n \to \infty$ uniformly for x_1 in M_1 and bounded $s, t > 0$. The equation (43) for $s, t > 0$ follows from this. Clearly, (43) is valid for those s, t for which the quantities in (43) are meaningful.

A theorem of van Kampen **(2)** implies that (41) has a unique solution locally. The conditions of van Kampen's theorem are that $f(p)$ is continuous; that (38) possesses a family of solutions $p = p(t; x_1)$, where $p(0; x_1) = x_1$ and $p(t; x_1)$ is defined on an open interval which can depend on x_1; that $p(t; x_1)$ is locally, uniformly Lipschitz continuous with respect to x_1; finally, that

(43) holds whenever the quantities in (43) are meaningful. The conclusion is that (41) has a unique solution locally.

This uniqueness assertion, together with (42), gives assertion (*). Hence the proof of (II) is complete.

Remark. If $\nu(c) = 0$ is permitted, (*) remains valid if "$d(p_1(t), p_2(t))$ is a decreasing" is replaced by "$d(p_1(t), p_2(t))$ is a non-increasing."

7. Proof of (II'). Let $x = x(t)$ be a solution of (21) defined at $t = S$. Write (21) as

$$x' = [f(t, x) - f(t, 0)] + [f(t, 0)].$$

Then an analogue of (20) is

$$x \cdot Gx' = \int_0^1 x \cdot GJ[s]x ds + x \cdot Gf(t, 0),$$

where $J[s] = J(t, xs)$. Thus, if $r^2 = x \cdot Gx$, it follows from $\alpha|x| < r < \beta|x|$, (32) and the monotony of x, that

$$r' < [-(r/\alpha)(\alpha/\beta^2)\lambda(r/\alpha) + |f(t, 0)|]\beta. \tag{47}$$

Assumption (33) implies that there exists a constant R such that

$$R > r(S) \text{ and } (R/\alpha)(\alpha/\beta^2)\lambda(R/\alpha) > |f(t, 0)| \text{ for } t > S.$$

Since $r(S) < R$, it is clear from (47) that $x = x(t)$ exists and that $r(t) < R$ for $t > S$.

References

1. W. Hahn, *Theorie und Anwendung der direkten Methode von Ljapunov* (Springer-Verlag, 1959).
2. E. R. van Kampen, *Remarks on systems of ordinary differential equations*, Amer. J. Math., *59* (1937), 144–152.
3. N. N. Krasovski, *Sufficient conditions for stability of solutions of a system of non-linear differential equations*, Doklady Akad. Nauk S.S.S.R. (N.S.), *98* (1954), 901–904 (Russian).
4. ——— *On stability in the large of the solutions of a non-linear system of differential equations*, Priklad. Math. i Mehanika, *18* (1954), 735–737 (Russian).
5. ——— *On stability with large initial disturbances*, Priklad. Math. i Mehanika, *21* (1957), 309–319 (Russian).
6. J. P. LaSalle, *Some extensions of Liapunov's second method*, Air Force Report AFOSR TN 60-22, 1960.
7. D. C. Lewis, *Differential equations referred to a variable metric*, Amer. J. Math., *73* (1951), 48–58.
8. L. Markus and H. Yamabe, *Global stability criteria for differential systems*, University of Minnesota report, mimeographed, 1960.
9. Z. Opial, *Sur la stabilité asymptotique des solutions d'un système d'équations différentielles*, Ann. Polonici Math., 7 (1960), 259–267.
10. A. Wintner, *Asymptotic integration constants*, Amer. J. Math., *68* (1946), 553–559, Appendix.

The Johns Hopkins University

ASYMPTOTIC SOLUTIONS OF EQUATIONS IN BANACH SPACE

C. A. SWANSON AND M. SCHULZER

1. Introduction. The equation $Px = y$ in Banach spaces has aroused considerable interest, particularly in view of the various situations in applied analysis which it encompasses, and consequently it has been the topic of numerous investigations **(2; 9; 10; 12)**. Detailed references may be found in **(10)**. The equation is of special interest because of its interpretation as an integral equation; and in turn, many problems related to differential equations can be reformulated as integral equations **(5; 7; 13)**.

Various iterative procedures are available **(10; 11; 12)** by which the existence and uniqueness of a solution x of such an equation can be established, and by which numerical estimates for the solution can be calculated. In any of these procedures, a sequence of elements x_n $(n = 0,1,2,\ldots)$ in the Banach space is constructed recursively, and is proved to converge in the Banach norm to an element x satisfying $Px = y$. The recursive sequences used have been modelled after various familiar ones. In particular, an iterative process modelled after Newton's method of solving real equations has been employed very successfully by Kantorovich **(10)** and others **(2; 16)**. Another recursive sequence, the analogue of that defined by an infinite continued fraction, has been studied recently by McFarland **(12)**. The most widely known iterative procedure is that based on the Liouville-Neumann sequence of successive approximations **(5; 11; 13)**.

The last of these, for example, can be used to prove that a contraction mapping T on a closed, bounded domain in the Banach space has a fixed point x in the domain **(1; 11)**. Therefore, under the assumption that the equation under consideration is equivalent to $Tx = x$ with T a contraction mapping, the existence and uniqueness of the solution follow from the fixed point theorem; and such results will be appropriate to the study of asymptotic properties of the solution.

In an investigation of the asymptotic behaviour of equations, one is interested in the variation of the elements y and the transformations P involved in the equations as a real variable λ (or more general variable) varies over an interval Λ. It is then pertinent to consider mappings (y) of Λ into the Banach space and mappings (P) of Λ into a suitable set of transformations on the space; and furthermore, in an asymptotic investigation, to study the behaviour of

Received February 3, 1960. This investigation was carried out while the second-named author held a National Research Council of Canada Scholarship.

these mappings as λ approaches a limit point, in general not in Λ. No essential features are lost by the assumptions that Λ is a positive interval $(0, \lambda_0]$ and that 0 is the limit point.

We shall first develop the notion of an asymptotically convergent sequence of mappings (§ 2), and from this, the notions of asymptotic equality and asymptotic summation of series. The main questions to be considered are the following: (1) If the quantities y and P involved in the equation $Px = y$ can be represented by asymptotically convergent series, can the solution be represented by such a series? (2) For a prescribed asymptotically convergent sequence of mappings (P_n), $n = 0, 1, 2, \ldots$, and for prescribed (y_n), does there exist a mapping (x) of Λ into the Banach space so that $\sum(P_n x)$ is asymptotically equal to $\sum(y_n)$? These questions are answered in § 5, in which the appropriate existence, uniqueness, and representation theorems are given. The approach taken here is similar to that employed by van der Corput **(14)** in connection with asymptotic solution of certain numerical equations.

We shall next mention a few examples, to which the subsequent theorems are applicable, obtained when the Banach space is specialized to one of the following: the space of real numbers; the finite dimensional Euclidean space V_n; the space of continuous functions on a closed, bounded interval; and the Lebesgue space L^p $(p > 1)$ **(13; 15)**.

In the space of real numbers x with norm defined by $||x|| = |x|$, the equation

$$y(\lambda) = \alpha x + \sum_{n=1}^{\infty} \alpha_n(\lambda, x) \qquad (\alpha \neq 0)$$

is to be considered, and the corresponding relation when asymptotic equality replaces equality. A specific example is the problem of finding a real number x with $|x| < 1$ so that for $|y| < 1$,

$$y \sim x + \sum_{n=1}^{\infty} (n+1)!\,(-\lambda)^n x^{n+1}.$$

In this example, formal substitution will lead to an asymptotic expansion

$$x \sim y + \sum_{n=1}^{\infty} \beta_n(\lambda) y^n$$

for the solution **(14)**. A different discussion of a similar problem has been given by de Bruijn **(6**, p. 25**)**.

As a second illustration, suppose the Banach space is specialized to the finite dimensional Euclidean space V_n. Each element x in V_n is a vector (σ_i) $(i = 1, 2, \ldots, n)$ with norm given by

$$||x|| = \left(\sum_{i=1}^{n} \sigma_i^2\right)^{\frac{1}{2}}.$$

In this context, one considers a system of n non-linear algebraic equations $y = A_0 x + Ex$, where y is a prescribed element of V_n, A_0 is a square matrix

of order n, and E is a transformation on V_n defined by $Ex = (E_k(\sigma_1, \sigma_2, \ldots, \sigma_n))$ $(k = 1, 2, \ldots, n)$. Under the assumption $\det A_0 \neq 0$, the linear system $y = A_0 x_0$ has a unique solution x_0. Under suitable additional hypotheses, Theorem 3 below guarantees that the non-linear system under consideration possesses a unique solution x such that $\|x - x_0\| \to 0$ as $\lambda \to 0$; and Theorem 5 gives an iterative procedure by which an asymptotic expansion can be generated.

We envisage that the most fruitful application will be to non-linear integral equations. The Banach space will be either the space C of all continuous functions over the closed, bounded interval under consideration, or the Lebesgue space L^p $(p > 1)$. The transformations A_0 and E will be regarded as linear and non-linear integral operators respectively. Consider the integral equation

$$x(s) = y(s) + \int_0^1 K(s,t)x(t)dt + \int_0^1 E_\lambda(s, t; x(t))dt \tag{1.1}$$

with $K(s, t)$ continuous on the closed unit square, and $y \in C(0, 1)$. This integral equation is of the form $x = y + Kx + Ex$, where K is the linear transformation from C into C and E is the non-linear transformation defined by (1.1). The first assumption to be made, of course, is that $(I - K)^{-1}$ exists, where I is the identity transformation, so that the linear integral equation approximating (1.1) for small λ will have a solution. The existence of this inverse transformation is implied by the condition $\|K\| < 1$ according to Banach's well-known theorem **(11; 13**, p. 151**)**. The analogue of the principal hypothesis (4.4) below is the hypothesis that $E_\lambda(s, t; u)$ satisfies a Lipschitz condition in its third argument on a suitable interval, uniformly for (s, t) on the unit square. Theorem 3 guarantees the existence of a unique solution $x = x(s, \lambda)$ of (1.1) with the property that $\|x - (I - K)^{-1}y\| \to 0$ as $\lambda \to 0$; and Theorem 5 shows how an asymptotic expansion of the solution can be generated by a recursive process. Similar statements can be made when the space C is replaced by the Hilbert space $L^2(0, 1)$.

2. Asymptotic convergence. A Banach space $\mathfrak{B}$ will be considered, and the Banach norm of an element $x \in \mathfrak{B}$ will be denoted as usual by $\|x\|$. The following notation will be used throughout: (i) λ denotes a positive real variable on an interval Λ_0: $0 < \lambda < \lambda_0$; (ii) ϕ denotes a function from Λ_0 into positive numbers; (iii) (x) denotes a mapping $\lambda \to x(\lambda)$ of Λ_0 into $\mathfrak{B}$; (iv) j, k, m, n denote non-negative integers; (v) $\alpha_0, \alpha_1, \ldots, \lambda_0, \lambda_1, \ldots,$ denote *fixed* positive numbers, that is positive numbers independent of λ.

Let ϕ_n $(n = 0, 1, 2, \ldots)$ be a single-valued function from Λ_0 into positive real numbers. The sequence $\{\phi_n\}$ is said to be an *asymptotic sequence* as $\lambda \to 0$ if $\phi_0(\lambda) = 1$ for all $\lambda \in \Lambda_0$, and $\phi_{n+1} = o(\phi_n)$ as $\lambda \to 0$ for each integer n **(8)**.

Let $\{\lambda_n\}$ be a non-increasing sequence of positive numbers, and for each integer n let Λ_n denote the interval $0 < \lambda < \lambda_n$.

Let $\{x_n(\lambda)\}$ $(n = 0, 1, 2, \ldots)$ be a sequence of elements in $\mathfrak{B}$, with $x_n(\lambda)$ uniquely defined for each $\lambda \in \Lambda_0$, and let (x_n) designate the mapping $\lambda \to x_n(\lambda)$ from Λ_0 into $\mathfrak{B}$. The sequence $\{(x_n)\}$ is said to *converge asymptotically* if there exists a single-valued mapping (x) of Λ_0 into $\mathfrak{B}$, an asymptotic sequence $\{\phi_n\}$, and a sequence of positive numbers α_n so that

$$(2.1) \qquad ||x(\lambda) - x_n(\lambda)|| < \alpha_n \phi_n(\lambda), \qquad \lambda \in \Lambda_n$$

for each integer n. In this event, (x) is referred to as an *asymptotic limit* of the sequence $\{(x_n)\}$. In particular, the sequence is said to converge asymptotically to zero when

$$(2.2) \qquad ||x_n(\lambda)|| < \alpha_n \phi_n(\lambda), \qquad \lambda \in \Lambda_n, \quad n = 0, 1, \ldots .$$

Our terminology follows that used by van der Corput in the asymptotic theory of numerical functions **(14)**.

An asymptotically convergent sequence need not converge in the ordinary sense (in the Banach norm) for any value of λ, as shown by the example $x_n(\lambda) = n!\lambda^n$ in the Banach space of real numbers.

Two mappings (x), (y) defined on Λ_0 are said to be asymptotically equal if for each integer n there exists a positive number α_n so that

$$(2.3) \qquad ||x(\lambda) - y(\lambda)|| < \alpha_n \phi_n(\lambda)$$

whenever $\lambda \in \Lambda_n$. In this event, we write $(x) \leftrightarrow (y)$. The relation $\leftrightarrow$ is evidently reflexive, symmetric, and transitive, and hence it is an equivalence relation among mappings. Each real asymptotic sequence $\{\phi_n\}$ induces such an equivalence relation, the sets of asymptotically equal mappings with respect to $\{\phi_n\}$ forming the equivalence classes.

If $\{(x_n)\}$ is an asymptotically convergent sequence of mappings, then the set of all asymptotic limits of the sequence is characterized by an equivalence class of asymptotically equal mappings. For let (x) be any asymptotic limit. Then if (y) is an asymptotic limit, it follows from (2.1) that

$$||x(\lambda) - y(\lambda)|| < ||x(\lambda) - x_n(\lambda)|| + ||y(\lambda) - x_n(\lambda)|| < \alpha_n \phi_n(\lambda) + \alpha_n' \phi_n(\lambda)$$

whenever $\lambda \in \Lambda_n$. Hence (2.3) holds and $(y) \leftrightarrow (x)$. Conversely, it is easy to see that if $(y) \leftrightarrow (x)$, then (y) is an asymptotic limit of the sequence.

Since $(x) \leftrightarrow (y)$ for any two asymptotic limits (x), (y) we shall say that the asymptotic limit of the sequence is *asymptotically unique.*

A formal series $\sum (x_n)$ is said to have an *asymptotic sum* (x) if (x) is an asymptotic limit of the sequence $\{(x_0 + x_1 + \ldots + x_{n-1})\}$ $(n = 1, 2, \ldots)$. This means that for each n there exists a positive number α_n and an interval Λ_n so that

$$(2.4) \qquad \left|\left|x(\lambda) - \sum_{j=0}^{n-1} x_j(\lambda)\right|\right| < \alpha_n \phi_n(\lambda), \qquad \lambda \in \Lambda_n.$$

When the asymptotic sum exists it is not unique, but it follows from the foregoing remarks that it is asymptotically unique. When (x) is an asymptotic

sum for $\sum(x_n)$, we say that the series is an asymptotic expansion for (x), and write $(x) \sim \sum(x_n)$.

The following theorem may be regarded as the basic theorem concerning asymptotic convergence. It states that an asymptotic sum of $\sum(x_n)$ always exists when $\{(x_n)\}$ converges asymptotically to zero. Results like this for numerical functions have been obtained by various authors **(3; 4; 8)**. The present proof is modelled after that of van der Corput **(14)**.

THEOREM 1. *A necessary and sufficient condition for a series of mappings $\sum(x_n)$ to have an asymptotic sum is that the sequence $\{(x_n)\}$ converge asymptotically to zero.*

Proof. If (x) is an asymptotic sum for $\sum(x_n)$, then (2.4) is valid for each integer n, and it is easily established from the Minkowski inequality and the order relation $\phi_{n+1} = o(\phi_n)$ $(\lambda \to 0)$ that (2.2) holds. Hence $\{(x_n)\}$ converges asymptotically to zero.

Conversely, if the sequence converges asymptotically to zero, then (2.2) holds for each integer n. Since $\phi_{n+1} = o(\phi_n)$ as $\lambda \to 0$, it follows that for each n there exist positive numbers α_n, λ_n with $\{\lambda_n\}$ non-increasing, so that

$$||x_{n+1}(\lambda)|| < \alpha_{n+1}\phi_{n+1}(\lambda) < \tfrac{1}{2}\alpha_n\phi_n(\lambda)$$

for $0 < \lambda < \lambda_{n+1}$, that is $\lambda \in \Lambda_{n+1}$. Hence

$$||x_{n+j}(\lambda)|| < (\tfrac{1}{2})^j\alpha_n\phi_n(\lambda) \qquad (j = 1, 2, \ldots) \tag{2.5}$$

for all $\lambda \in \Lambda_{n+j}$.

If λ_n tends to a positive limit λ^* as $n \to \infty$, it follows from (2.5) that

$$\left\{\sum_{j=0}^{n} x_j(\lambda)\right\}$$

is a Cauchy sequence for each λ satisfying $0 < \lambda < \lambda^*$. Hence this sequence converges in the Banach norm to an element $x(\lambda)$ because of the completeness of $\mathfrak{B}$. It can then be verified that $x(\lambda)$ satisfies (2.4), and consequently (x): $\lambda \to x(\lambda)$ is an asymptotic sum.

The situation of real interest, however, is that in which $\sum x_n(\lambda)$ does *not* have an ordinary sum for any positive value of λ. Suppose then that $\lambda_n \to 0$ as $n \to \infty$. For each value of λ, let $H = H(\lambda)$ be the largest integer such that $\lambda_H > \lambda$. Then if $\lambda \in \Lambda_n$, it follows that $H(\lambda) > n$. We assert that (x) given by

$$(x) = (x_0 + x_1 + \ldots + x_H)$$

is an asymptotic sum for $\sum(x_n)$. In fact, for all $\lambda \in \Lambda_n$, $H(\lambda)$ has been chosen so that $\lambda \in \Lambda_H$, which implies that $\lambda \in \Lambda_{n+j}$ $(j = 0, 1, \ldots, H - n)$. Then by (2.5)

$$\left|\left|x(\lambda) - \sum_{j=0}^{n-1} x_j(\lambda)\right|\right| < \sum_{j=n}^{H} ||x_j(\lambda)|| < 2\alpha_n\phi_n(\lambda), \tag{2.6}$$

and hence (x) is an asymptotic sum.

3. Transformations on the Banach space. A *transformation* E defined on a closed domain $\mathfrak{D}$ in $\mathfrak{B}$ is a single-valued mapping from $\mathfrak{D}$ into $\mathfrak{B}$. Transformations are not necessarily additive, nor are they necessarily even defined on the whole space $\mathfrak{B}$. For each $\lambda \in \Lambda_0$, let $E(\lambda)$ be a uniquely defined transformation on $\mathfrak{D}$, and let (E) be the mapping $\lambda \to E(\lambda)$. We shall say that (E) is in the class Lip $(\mathfrak{D}, \phi)$ whenever there exists a fixed, positive number α and a bounded positive function ϕ on Λ_0 so that

$$(3.1) \qquad ||E(\lambda)x - E(\lambda)y|| \leqslant \alpha\phi(\lambda)||x - y||$$

for all pairs of elements x, y in $\mathfrak{D}$, and for all $\lambda \in \Lambda_0$. When $\alpha\phi(\lambda) < 1$, a transformation $E(\lambda)$ from $\mathfrak{D}$ into itself satisfying (3.1) is a contraction mapping, and therefore has a fixed point in $\mathfrak{D}$ **(11)**.

Sums and products of transformations are defined as in the linear case **(13**; **15)**. Thus, if E, F are transformations on $\mathfrak{D}$, then $(E + F)x = Ex + Fx$, $x \in \mathfrak{D}$; and $(EF)x = E(Fx)$, $x \in \mathfrak{D} \cap \mathfrak{R}$, where $\mathfrak{R}$ is the range of F.

It is convenient to introduce the symbol $||E||$ to denote the supremum of $||Ex - Ey||/||x - y||$ over all x, y with $x \neq y$. Then (3.1) may be rewritten in the form

$$(3.2) \qquad ||E(\lambda)|| \leqslant \alpha\phi(\lambda), \qquad \lambda \in \Lambda_0.$$

For each integer n, let (A_n) denote a mapping $\lambda \to A_n(\lambda)$ of a positive interval Λ_n into the set of transformations on $\mathfrak{D}$. The sequence $\{(A_n)\}$ is said to *converge asymptotically* to (A) on $\mathfrak{D}$ whenever there exists an asymptotic sequence $\{\phi_n\}$ so that $(A) \in \mathrm{Lip}(\mathfrak{D}, \phi_0)$ and $(A - A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ for each integer n. In this event,

$$(3.3) \qquad ||A(\lambda) - A_n(\lambda)|| \leqslant \alpha_n\phi_n(\lambda), \qquad \lambda \in \Lambda_n.$$

(A) will be called an *asymptotic limit* of the sequence $\{(A_n)\}$.

A series $\sum (A_n)$ is said to have an *asymptotic sum* (A) if (A) is an asymptotic limit of the sequence $\{(A_0 + A_1 + \ldots + A_{n-1})\}$. This means that there exists a positive number α_n so that

$$(3.4) \qquad \left|\left|A(\lambda) - \sum_{j=0}^{n-1} A_j(\lambda)\right|\right| \leqslant \alpha_n\phi_n(\lambda), \qquad \lambda \in \Lambda_n.$$

The following analogue of Theorem 1 is valid for transformations.

THEOREM 2. *A necessary and sufficient condition for the series* $\sum_{n=m}^{\infty}(A_n)$ *to have an asymptotic sum in the class* $\mathrm{Lip}(\mathfrak{D}, \phi_m)$ *is that* $(A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ *for each integer* $n \geqslant m$.

The proof parallels that of Theorem 1. To establish the sufficiency, we choose the integer $H(\lambda)$ as in Theorem 1 and define (A) by

$$A(\lambda) = A_m(\lambda) + A_{m+1}(\lambda) + \ldots + A_H(\lambda).$$

Then the mapping $(A): \lambda \to A(\lambda)$ will be an asymptotic sum; in fact, from $(A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ it follows that $||A_{n+1}(\lambda)|| < \frac{1}{2}\alpha_n\phi_n(\lambda)$, and hence that

$$\left|\left|A(\lambda) - \sum_{j=m}^{n-1} A_j(\lambda)\right|\right| < 2\alpha_n\phi_n(\lambda), \qquad \lambda \in \Lambda_n$$

for each integer $n > m + 1$. In particular, $(A) \in \mathrm{Lip}(\mathfrak{D}, \phi_m)$.

4. Equations in Banach space. For a prescribed element $y(\lambda)$ in $\mathfrak{B}$ it will be our purpose to obtain information concerning the solution x of the equation

$$P(\lambda)x = y(\lambda). \tag{4.1}$$

The element $y = y(\lambda)$ is supposed to be uniquely defined for each λ in a positive interval Λ_0. The mapping $(y): \lambda \to y(\lambda)$ is supposed to possess an asymptotic expansion

$$(y) \sim \sum_{n=0}^{\infty} (y_n) \qquad (\lambda \to 0) \tag{4.2}$$

in which y_0 is a fixed element of $\mathfrak{B}$. According to (2.4), this means in particular that $||y(\lambda) - y_0|| \to 0$ as $\lambda \to 0$.

It will be assumed that the transformation $P(\lambda)$ in (4.1) is uniquely defined on some fixed domain $\mathfrak{D}' \subset \mathfrak{B}$ for each $\lambda \in \Lambda_0$, and that $P(\lambda)$ has the decomposition

$$P(\lambda) = A_0 + E(\lambda), \qquad \lambda \in \Lambda_0 \tag{4.3}$$

valid on the entire domain of definition of $P(\lambda)$. In (4.3) A_0 is a fixed linear transformation with bounded inverse A_0^{-1}, and $E(\lambda)$ is a suitable contraction mapping, to be made precise presently.

For a fixed positive number η, let $\mathfrak{D}$ denote the closed sphere $\{x \in \mathfrak{B}: ||x - A_0^{-1}y_0|| < \eta\}$. The following assumptions will be made.

(i) $A_0^{-1}y_0 \in \mathfrak{D}'$ *and* η *is chosen small enough so that* $\mathfrak{D}$ *is a subset of* $\mathfrak{D}'$.

(ii) *The mapping* $(E): \lambda \to E(\lambda)$ *has the property that*

$$(E) \in \mathrm{Lip}(\mathfrak{D}, \phi_1) \qquad (\phi_1 = o(1) \quad as \quad \lambda \to 0). \tag{4.4}$$

(iii) *For any element* $z \in \mathfrak{D}$

$$||E(\lambda)z|| = o(1) \qquad as \quad \lambda \to 0. \tag{4.5}$$

The assumptions (4.3) and (4.4) together constitute a statement of the approximate linearity of $P(\lambda)$ in the neighbourhood of $\lambda = 0$; there exists a linear transformation A_0 and a positive interval Λ_1 so that $||P(\lambda) - A_0|| < \alpha_1\phi_1(\lambda)$ whenever $\lambda \in \Lambda_1$, and $\phi_1(\lambda) \to 0$ as $\lambda \to 0$.

Now we shall establish an existence and uniqueness theorem appropriate to the study of asymptotic properties of the solution, by appealing to the theorem that every contraction mapping on $\mathfrak{D}$ has a fixed point in $\mathfrak{D}$.

THEOREM 3. *Under the assumptions* (4.3), (4.4), (4.5) *there exists a positive interval* Λ *so that the equation* $P(\lambda)x = y(\lambda)$ *has a unique solution* $x(\lambda) \in \mathfrak{D}$ *for each* $\lambda \in \Lambda$. *Furthermore* $||x(\lambda) - A_0^{-1}y(\lambda)|| \to 0$ *as* $\lambda \to 0$.

Proof. On account of (4.3), equation (4.1) is equivalent to

$$x = v(\lambda) + F(\lambda)x, \tag{4.6}$$

where $v(\lambda) = A_0^{-1}y(\lambda)$ and $F(\lambda) = -A_0^{-1}E(\lambda)$. This equation has the form $x = T(\lambda)x$. We shall demonstrate that there exists a positive interval Λ so that $T(\lambda)$ maps $\mathfrak{D}$ into $\mathfrak{D}$ whenever $\lambda \in \Lambda$. In fact, if $x \in \mathfrak{D}$, that is $||x - v_0|| < \eta$ where $v_0 = A_0^{-1}y_0$, then

$$||T(\lambda)x - v_0|| < ||v(\lambda) - v_0|| + ||F(\lambda)x||.$$

However,

$$||v(\lambda) - v_0|| < ||A_0^{-1}||\,||y(\lambda) - y_0|| \to 0 \quad \text{as} \quad \lambda \to 0 \text{ by } (4.2),$$

and

$$||F(\lambda)x|| < ||A_0^{-1}||\,||E(\lambda)x|| \to 0 \quad \text{as} \quad \lambda \to 0 \quad \text{by} \quad (4.5).$$

Therefore there exists a positive interval Λ so that $||T(\lambda)x - v_0|| < \eta$ whenever $\lambda \in \Lambda$, and hence $T(\lambda)x \in \mathfrak{D}$ whenever $\lambda \in \Lambda$.

Clearly $(F) \in \mathrm{Lip}(\mathfrak{D}, \phi_1)$ since $(E) \in \mathrm{Lip}(\mathfrak{D}, \phi_1)$. Then

$$||T(\lambda)x - T(\lambda)y|| = ||F(\lambda)x - F(\lambda)y|| < \alpha\phi_1(\lambda)||x - y||$$

for all $x, y \in \mathfrak{D}$ and all λ in some positive interval. Since $\phi_1 = o(1)$ as $\lambda \to 0$, there exists a positive interval Λ' so that $||T(\lambda)x - T(\lambda)y|| < \frac{1}{2}||x - y||$ whenever $\lambda \in \Lambda'$; $x, y \in \mathfrak{D}$. We may assume that $\Lambda' = \Lambda$. Then $T(\lambda)$ is a contraction mapping on $\mathfrak{D}$, a closed sphere in the complete space $\mathfrak{B}$, for all $\lambda \in \Lambda$. Hence $T(\lambda)$ has a fixed point $x = x(\lambda) \in \mathfrak{D}$ (**11**), that is $x(\lambda)$ satisfies (4.6) and hence (4.1).

Finally,

$$||x(\lambda) - A_0^{-1}y(\lambda)|| = ||x(\lambda) - v(\lambda)|| = ||Fx(\lambda)|| \to 0$$

as $\lambda \to 0$ by (4.5).

For the validity of this theorem, the assumption (4.4) is needed to ensure that the mapping T be a contraction mapping. It is well known that a stronger condition than continuity of the transformation E is required to imply uniqueness of the solution: such a condition is the Lipschitz condition (4.4). Counter-examples can be easily supplied when (4.1) is interpreted as an integral equation (for example, of the type arising from an initial value problem for a differential equation (**5**, chapter I)).

Assumption (4.5) is needed so that T maps $\mathfrak{D}$ into itself. A simple counter-example in the Banach space of real numbers to show that Theorem 3 is false without such an assumption is provided by the real equation $x = 2 + (\lambda x^{\frac{1}{2}} - 3)$ with $Ex = \lambda x^{\frac{1}{2}} - 3$, which does not have a solution in the space.

5. Asymptotic solution of equations. Consider now a mapping (A_n): $\lambda \to A_n(\lambda)$ in the class $\mathrm{Lip}(\mathfrak{D}, \phi_n)$ for each integer $n = 0, 1, 2, \ldots$, where $\mathfrak{D}$ is the closed sphere defined in the previous section. It will be assumed that A_0 is a fixed linear transformation on $\mathfrak{B}$ with a bounded inverse. We seek a mapping $(x): \lambda \to x(\lambda) \in \mathfrak{D}$ for which a prescribed (y) is an asymptotic sum for the series $\sum (A_n x)$. Such an x will be called an *asymptotic solution* of the relation $\sum (A_n x) \sim (y)$. Next, a theorem will be derived concerning asymptotic solutions, under the following assumptions: (i) $(A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ $(n = 1, 2, \ldots)$; (ii) *For any element* $z \in \mathfrak{D}$

$$\|A_n z\| < \alpha_n \phi_n(\lambda) \qquad (\lambda \in \Lambda_n, \qquad \alpha_n > 0). \tag{5.1}$$

THEOREM 4. *Under these assumptions, there exists an asymptotically unique solution* $(x): \lambda \to x(\lambda) \in \mathfrak{D}$ *of the relation*

$$\sum_{n=0}^{\infty} (A_n x) \sim (y).$$

Proof. Since $(A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ for each integer n, it follows from Theorem 2 that there exists an asymptotic sum (P) of

$$\sum_{n=0}^{\infty} (A_n)$$

defined on $\mathfrak{D}$; and in fact (P) is given by

$$P(\lambda) = \sum_{n=0}^{H(\lambda)} A_n(\lambda),$$

where $H(\lambda)$ is a suitable integer depending on λ. It follows in particular that the mapping $(E) = (P - A_0)$ is in the class $\mathrm{Lip}(\mathfrak{D}, \phi_1)$, which is Assumption (4.4) of Theorem 3. Since the sequence $\{(A_n x)\}$ converges asymptotically to zero for any $x \in \mathfrak{D}$ by (5.1) it follows from Theorem 1 that (Px) is an asymptotic sum for $\sum (A_n x)$, $x \in \mathfrak{D}$. In particular, $\|E(\lambda)x\| = \|[P(\lambda) - A_0]x\| < \alpha_1\phi_1(\lambda)$ $(\lambda \in \Lambda_1)$, which is the content of Assumption (4.5) of Theorem 3. Therefore the assumptions of the present theorem imply those of Theorem 3, and there exists an element $x(\lambda) \in \mathfrak{D}$ satisfying $P(\lambda)x(\lambda) = y(\lambda)$ $(\lambda \in \Lambda)$. Then for $\lambda \in \Lambda_n$, we obtain from (5.1)

$$\left\| y(\lambda) - \sum_{j=0}^{n-1} A_j(\lambda)x(\lambda) \right\| < \sum_{j=n}^{H} \|A_j(\lambda)x(\lambda)\| < 2\alpha_n\phi_n(\lambda) \tag{5.2}$$

by the same reasoning which led to (2.5) and (2.6), and hence (y) is an asymptotic sum for $\sum (A_n x)$.

To show that (x) is asymptotically unique, let (u) be any other asymptotic solution. Then for each integer n,

$$\|A_0(x-u)\| - \left\|\sum_{j=1}^{n-1} (A_jx - A_ju)\right\| < \left\|\sum_{j=0}^{n-1} A_jx - A_ju\right\|$$
$$< \left\|y - \sum_{j=0}^{n-1} A_jx\right\| + \left\|y - \sum_{j=0}^{n-1} A_ju\right\|,$$

and according to (5.2) there is a positive number β_n ($n = 1, 2, \ldots$) so that

$$\|A_0(x-u)\| - \left\|\sum_{j=1}^{n} (A_jx - A_ju)\right\| < \beta_n\phi_n(\lambda), \qquad \lambda \in \Lambda_n. \tag{5.3}$$

By hypothesis, there exists a positive number α so that $\|A_0^{-1}\| < \alpha$, and hence

$$\|x-u\| < \|A_0^{-1}\|\,\|A_0(x-u)\| < \alpha\|A_0(x-u)\|. \tag{5.4}$$

Since $(A_n) \in \mathrm{Lip}(\mathfrak{D}, \phi_n)$ and $\{\phi_n\}$ is an asymptotic sequence it follows that there exists a sequence of positive intervals Λ_n' so that

$$\left\|\sum_{j=1}^{n-1} A_jx - A_ju\right\| < \sum_{j=1}^{n-1} \alpha_j\phi_j(\lambda)\|x-u\| < 2\alpha_1\phi_1(\lambda)\|x-u\|$$

whenever $\lambda \in \Lambda_n'$. Since $\phi_1(\lambda) \to 0$ as $\lambda \to 0$, there is a positive interval Λ so that $2\alpha_1\phi_1(\lambda) < \frac{1}{2}\alpha^{-1}$ whenever $\lambda \in \Lambda$. We may assume that $\Lambda_n' \subseteq \Lambda$, and hence

$$\left\|\sum_{j=1}^{n} A_jx - A_ju\right\| < \tfrac{1}{2}\alpha^{-1}\|x-u\|, \qquad \lambda \in \Lambda_n'. \tag{5.5}$$

Then (5.3), (5.4), and (5.5) together establish that

$$\tfrac{1}{2}\alpha^{-1}\|x-u\| < \beta_n\phi_n(\lambda)$$

whenever λ is in the smaller of the positive intervals Λ_n', Λ_n. Hence (x) is asymptotically unique.

In our final theorem, we shall derive an asymptotic expansion for the asymptotically unique solution (x) of $\sum(A_jx) \sim (y)$ given in Theorem 4. Suppose that (y) has the asymptotic expansion (4.2). Suppose also that (P) is an asymptotic sum for the series $\sum(A_n)$, as in Theorem 4. Then the mappings (v) and (F) defined in (4.6) will be asymptotic sums for corresponding series $\sum(v_n)$, $\sum(F_n)$, that is

$$\begin{aligned} &(v) \sim \textstyle\sum (v_n) \quad \text{where} \quad v_n = A_0^{-1}y_n \\ &(F) \sim \textstyle\sum (F_n) \quad \text{where} \quad F_n = -A_0^{-1}A_n. \end{aligned} \tag{5.6}$$

For the solution $(x): \lambda \to x(\lambda)$, it follows from Theorem 4 that $x(\lambda)$ satisfies equation (4.1), and hence satisfies (4.6). An asymptotic expansion for (x) will be obtained in a natural way from a sequence of successive approximations to the solution of (4.6), defined in terms of the quantities v_n, F_n.

The additional hypothesis will be made that the sequence $\{\phi_n\}$ has the multiplicative property

$$(5.7) \qquad \sum_{j=1}^{n} \phi_j(\lambda)\phi_{n-j}(\lambda) < \gamma_n\phi_n(\lambda) \qquad (\lambda \in \Lambda_0; n = 1, 2, \ldots)$$

where γ_n is a fixed positive number for each integer n.

THEOREM 5. *Under the hypotheses of Theorem 4, the asymptotically unique solution (x) of the relation $\sum(A_n x) \sim (y)$ is an asymptotic limit of the sequence $\{(x_n)\}$ defined by*

$$(5.8) \qquad x_0 = v_0; \qquad x_n = \sum_{j=0}^{n} v_j + \sum_{j=1}^{n} F_j x_{n-j} \qquad (n = 1, 2, \ldots).$$

An equivalent conclusion is that (x) has the asymptotic expansion

$$\sum(x_n - x_{n-1}) \qquad (\text{with} \quad x_{-1} = 0).$$

Proof. It is enough to show that there exists a sequence of positive numbers β_n and a sequence of intervals Λ_n so that

$$(5.9) \quad \|x(\lambda) - x_{n-1}(\lambda)\| < \beta_n\phi_n(\lambda) \qquad (\lambda \in \Lambda_n; \quad n = 1, 2, \ldots).$$

This will be proved by mathematical induction on n. First, it is easily seen from (5.2) and (5.6) that the proposition is true for $n = 1$. Under the hypothesis that it is true for all integers $j < n - 1$, we shall show that it is true for n. Since $x = v + Fx$, it follows that

$$(5.10) \quad \|x - x_n\| < \left\|v - \sum_{j=0}^{n} v_j\right\| + \left\|Fx - \sum_{j=1}^{n} F_j x\right\| + \left\|\sum_{j=1}^{n} F_j x - F_j x_{n-j}\right\|.$$

On account of the hypotheses (5.6) there exists a positive number α_{n+1} so that each of the first two terms on the right side of (5.10) is bounded above by $\alpha_{n+1}\phi_{n+1}(\lambda)$ for all λ in a positive interval Λ_{n+1}. The inductive proof of (5.9) will then be finished if it can be shown that the third term also is of order ϕ_{n+1}. To see this, observe that

$$\begin{aligned}\left\|\sum_{j=1}^{n} F_j x - F_j x_{n-j}\right\| &< \sum_{j=1}^{n} \|F_j\|\,\|x - x_{n-j}\| \\ &< \|A_0^{-1}\| \sum_{j=1}^{n} \|A_j\|\,\|x - x_{n-j}\| \\ &< \alpha \sum_{j=1}^{n} \alpha_j\phi_j(\lambda)\beta_{n-j+1}\phi_{n-j+1}(\lambda),\end{aligned}$$

where use has been made of the inductive hypothesis (5.9) and the hypothesis $(A_j) \in \mathrm{Lip}(\mathfrak{D}, \phi_j)$ at the last step. Let

$$\delta_n = \max_j \alpha_j\beta_{n-j+1} \qquad (1 < j < n).$$

Then, since $\{\phi_n\}$ has the multiplicative property (5.7),

$$\left|\left| \sum_{j=1}^{n} F_j x - F_j x_{n-j} \right|\right| < \alpha \delta_n \gamma_{n+1} \phi_{n+1}(\lambda).$$

Hence (5.9) is valid for each integer n, and the theorem is proved.

References

1. S. Banach, *Théorie des opérations linéaires*, Monografje Matematyczne, I (Warsaw, 1932).
2. R. G. Bartle, *Newton's method in Banach spaces*, Proc. Amer. Math. Soc., *6* (1955), 827–831.
3. E. Borel, *Sur quelques points de la théorie des fonctions*, Thèse, Annales de l'Ecole Normale (1895).
4. T. G. T. Carleman, *Les fonctions quasi-analytiques* (Paris, 1926), chapter 5.
5. E. A. Coddington and N. Levinson, *Theory of ordinary differential equations* (McGraw-Hill, New York, 1955).
6. N. G. de Bruijn, *Asymptotic methods in analysis* (P. Noordhoff, Groningen, Netherlands, 1958).
7. G. F. D. Duff, *Partial differential equations* (University of Toronto Press, Toronto, 1956).
8. A. Erdélyi, *Asymptotic expansions* (Dover, New York, 1956).
9. Yu. Ya. Kaazik and E. E. Tamme, *On a method of approximate solution of functional equations*, Dokl. Akad. Nauk SSSR (N.S.), *101* (1955), 981–984 (Russian).
10. L. V. Kantorovich, *Functional analysis and applied mathematics*, Translated by C. D. Benster, National Bureau of Standards (U.C.L.A., 1952).
11. A. N. Kolmogoroff and S. V. Fomin, *Elements of the theory of functions and functional analysis I* (Graylock, Rochester, N.Y., 1957).
12. J. E. McFarland, *An iterative solution of the quadratic equation in Banach space*, Proc. Amer. Math. Soc., *9* (1958), 824–830.
13. F. Riesz and B. Sz.-Nagy, *Functional analysis* (Frederick Ungar, New York, 1955).
14. J. G. van der Corput, *Asymptotic expansions I*, Technical report I, Contract AF-18(600)-958 (University of California, Berkeley, 1954).
15. A. C. Zaanen, *Linear analysis* (Interscience, New York, 1953).
16. D. M. Zagadskii, *An analogue of Newton's method for non-linear integral equations*, Dokl. Akad. Nauk SSSR (N.S.), *59* (1948), 1041–1044 (Russian).

University of British Columbia

NORMAL OPERATORS ON THE BANACH SPACE $L^p(-\infty,\infty)$. PART I

GREGERS L. KRABBE

1. Introduction. Let $\mathfrak{B}\bar{R}^2$ be the Boolean algebra of all finite unions of subcells of the plane. Denote by $\mathscr{E}_p$ the algebra of all linear bounded transformations of $L^p(-\infty, \infty)$ into itself. Suppose for a moment that $p = 2$, and let $\mathscr{R}_p$ be an involutive abelian subalgebra of $\mathscr{E}_p$: if $\mathscr{R}_p$ is also a Banach space and if $T_p \in \mathscr{R}_p$, then:

(i) *The family of all homomorphic mappings of* $\mathfrak{B}\bar{R}^2$ *into the algebra* $\mathscr{R}_p$ *contains a member* $E_p{}^T$ *such that*

$$T_p = \int \lambda \cdot E_p^T(d\lambda). \tag{1}$$

Suppose, henceforth, that $1 < p < \infty$. The main result of this article (Theorem 6.14) shows that property (i) remains valid for a suitable algebra $\mathscr{R}_p$.

Let $\mathfrak{D}$ be the class of all bounded functions whose real and imaginary parts are piecewise monotone. In § 2 will be defined an isomorphism $f \to [\mathsf{A}f]_p$ whose domain includes $\mathfrak{D}$ and whose range $(t)_p$ is a normed involutive abelian subalgebra of $\mathscr{E}_p$. Theorem 6.14 will show that a member T_p of $(t)_p$ has the property (i) whenever $T_p = [\mathsf{A}f]_p$ for some f in $\mathfrak{D}$. The relation (1) involves a Riemann–Stieltjes integral defined in the strong operator-topology of $\mathscr{E}_p$ (see 6.11). The set-function $E_p{}^T$ need not be countably additive: we do not restrict ourselves to "*spectral resolutions*" in the sense of Dunford **(1)**. The values of $E_p{}^T$ are self-adjoint **(4**, p. 22**)**, idempotent members of $(t)_p$.

It is easily seen that the Hilbert transformation and the Dirichlet operators all have the property (i). For less trivial examples, let $\mathscr{M}^1$ be the set of all bounded Radon measures; if $A \in \mathscr{M}^1$, then the convolution operator $A*_p$ is defined as the mapping $x \to A*x$ of $L^p(-\infty, \infty)$ into itself. In the special case $A \in L^1(-\infty, \infty)$, the operator $A*_p$ is defined for all x in $L^p(-\infty, \infty)$ by the relation

$$A*_p x(\theta) = \int_{-\infty}^{\infty} A(\theta - \beta)x(\beta)d\beta.$$

In case the Fourier transform of A belongs to $\mathfrak{D}$, then the operator $T_p = A*_p$ satisfies property (i). Consequently, all the classical convolution operators (Picard, Poisson, Weierstrass, Stieltjes, Fejér, etc.) have property (i). Explicit determination of $E_p{}^T$ is readily inferred from § 6; in the case $p = 2$ our results

Received February 24, 1960. This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under Contract No. AF 49(638)-505.

coincide with the ones given by Dunford (**2**, p. 63) for operators T of this type. The completion of the algebra $\{A_{*p}: A \in \mathcal{M}^1\}$ is an (A^*)-subalgebra of $(t)_p$ (see 3.2).

Let ∇_p be the operator defined by the relation

$$\nabla_p x = \frac{i}{2\pi} \text{ (derivative of x)}$$

for all x in a suitable subset of $L^p(-\infty, \infty)$; this unbounded operator also has the property (i). Although details regarding such operators will be reserved for a subsequent article (see 7.0), it may be pertinent to remark here that a relation of the type

$$T_p = \int_{-\infty}^{\infty} f(\theta) E_p^{\nabla} (d\theta)$$

holds for any T_p in $(t)_p$ such that $T_p = [\wedge f]_p$ for some function f of locally bounded variation. For example, take α in $(-\infty, \infty)$, and let R_p be the translator defined by $R_p x(\theta) = x(\theta - \alpha)$ for all x in $L^p(-\infty, \infty)$; then

$$R_p = \int_{-\infty}^{\infty} e^{2\pi i\alpha\theta} E_p^{\nabla} (d\theta).$$

2. The basic function-algebra. Let $\mathfrak{F}_+$ denote the set of all complex-valued measurable functions defined on $(-\infty, \infty)$. Note that $\mathfrak{F}_+$ is an algebra with multiplication $f \cdot g = \{\theta \to f(\theta) g(\theta)\}$. The customary identification of equivalent functions is implied henceforth.

Let L^+ be the intersection of the family $\{L^p(-\infty, \infty): 1 < p < \infty\}$. The Fourier transform Ψz of a function z in L^+ is defined as the function f such that $||f - f_n||_2 \to 0$, where $n \to \infty$ and

$$f_n(\theta) = \int_{-n}^{n} e^{2\pi i\theta\beta} z(\beta) d\beta \qquad (-\infty < \theta < \infty).$$

We denote by (t^+) the set of all linear mappings of L^+ into itself. If $T \in (t^+)$, then

$$|T|_p = \text{su p}\{||Tx||_p : x \in L^+ \text{ and } ||x||_p < 1\}.$$

Let $\mathcal{E}$ denote the set of all T in (t^+) such that $|T|_p \neq \infty$ whenever $1 < p < \infty$. If $G \in \mathfrak{F}_+$, then $t(g)$ is defined as the set of all T in $\mathcal{E}$ such that

$$\Psi(Tx) = g \cdot \Psi x \qquad \text{for all } x \text{ in } L^+. \tag{2}$$

2.1. *Definition.* Let $\mathfrak{F}$ denote the algebra of all bounded members of $\mathfrak{F}_+$. Our basic operator-algebra is the set

$$(t) = \bigcup \{t(g): g \in \mathfrak{F}\}.$$

If $T \in (t)$, then $\mathbf{v}T$ will denote the unique g in $\mathfrak{F}$ such that $T \in t(g)$. The set $\{\mathbf{v}T: T \in (t)\}$ is denoted by $\mathfrak{F}_{\mathbf{v}}$.

2.2. *Remarks.* The definition of $\mathbf{v}T$ is justified by the fact that $g = \mathbf{O}$ whenever $g \cdot \Psi x = \mathbf{O}$ for all x in L^+. Note that $\mathfrak{F}_{\mathbf{v}}$ is the set of all g in $\mathfrak{F}$ such that $\emptyset \neq t(g)$. It is easily checked that (t) is an abelian subalgebra of $\mathscr{E}$ and that $\{T \to \mathbf{v}T\}$ maps (t) isomorphically onto $\mathfrak{F}_{\mathbf{v}}$; in particular

$$(3) \qquad \mathbf{v}(T^{(0)}T^{(1)}) = (\mathbf{v}T^{(0)}) \cdot (\mathbf{v}T^{(1)}) \qquad \text{when} \quad T^{(n)} \in (t).$$

2.3. *Notation.* If $x \in L^p(-\infty, \infty)$, let $x^{\cdot} = \{\theta \to x(-\theta)\}$, while $\bar{x} = \{\theta \to \overline{x(\theta)}\}$ and $\sim x = \{\theta \to \overline{x(-\theta)}\}$.

2.4. *Remarks.* If $T \in \mathscr{E}$ we define $\sim T$ as the operator $\{x \in L^+ \to \sim T \sim x\}$; observe that $|T|_p = |\sim T|_p$ (this follows from $\|x\|_p = \|\sim x\|_p$). If $T \in t(g)$, then it is easily checked that $\sim T \in t(\bar{g})$. Therefore, the mapping $\{T \to \sim T\}$ of (t) into itself is an *involution* (**10**, p. 108).

2.5. The following terminology is found in Hille (**4**, p. 22): a member T of $\mathscr{E}$ is "*self-adjoint*" if $T = \sim T$. It is clear that T will be self-adjoint if and only if the function $\mathbf{v}T$ is real-valued.

3. The basic operator-algebra. From now on, p is a fixed number $(1 < p < \infty)$. Let $\mathscr{E}_p$ denote the Banach space of all bounded linear transformations of $L^p(-\infty, \infty)$ into itself. Since L^+ is dense in $L^p(-\infty, \infty)$, each T in $\mathscr{E}$ has a unique, continuous extension T_p in $\mathscr{E}_p$. Consequently, the algebra (t) is isomorphic to $(t)_p = \{T_p: T \in (t)\}$ under the mapping $\{T \to T_p\}$. Note that $|T_p|_p = |T|_p$. From 2.4 it follows that $(t)_p$ is a normed involutive subalgebra (**10**, p. 110) of $\mathscr{E}_p$ in the sense that $|T_p|_p = |\sim T_p|_p$. Note further that $(t)_p$ contains the identity operator $\mathbf{I}_p = \{x \in L^p(-\infty, \infty) \to x\}$, and the completion $(t)_p^*$ of $(t)_p$ is a (*)-algebra in the sense of (**4**, p. 22). The title of this article was suggested by the fact that all members of $(t)_p$ are "normal" (**4**, p. 22).

3.1. *Application.* Let $\mathscr{M}^1$ be the algebra of all bounded Radon measures on $(-\infty, \infty)$. If $A \in \mathscr{M}^1$, then A_* is defined as the mapping $\{x \to A * x\}$ of L^+ into itself (where $A*x$ = convolution of A and x; see (**9**)). In 3.2 it will be shown that the completion of $\mathscr{A}_p = \{A_{*p}: A \in \mathscr{M}^1\}$ is an (A^*)-subalgebra of $(t)_p^*$ (see (**4**, Definition 1.15.3)). It is known that $A_* \in \mathscr{E}$. If $\Psi(dA)$ is the function g defined by

$$g(\theta) = \int_{-\infty}^{\infty} e^{2\pi i\theta\beta} dA(\beta) \qquad (-\infty < \theta < \infty),$$

then $\Psi(A*x) = \Psi(dA) \cdot \Psi x$ (this can be seen from (**9**, p. 133, (II)), where $\Psi(dA)$ is denoted (YA)). But $\Psi(dA) \in \mathfrak{F}$, whence $A_* \in (t)$ and $\mathbf{v}A_* = \Psi(dA)$. Consequently:

3.2. *If* $T_p = A_{*p}$ *and* $A \in \mathscr{M}^1$, *then* $T_p \in (t)_p$ *and* $\mathbf{v}T = \Psi(dA)$. Thus $\mathscr{A}_p \subset (t)_p$. To show that the completion of $\mathscr{A}_p$ is an (A^*)-algebra, suppose that $T_p = A_{*p}$ is self-adjoint; from 2.5, 3.2, and (**9**, (i)) it follows that the spectrum of T_p is real.

3.3. *Definitions.* If $f \in \mathfrak{F}_\mathbf{v}$, we denote by $[\wedge f]$ the inverse image of f under the mapping $\{T \to \mathbf{v}T\}$; in other words, $[\wedge f]$ is the member T of (t) such that $f = \mathbf{v}T$. If $p' = p/(p-1)$ and $L^p = L^p(-\infty, \infty)$, then

$$\langle x, y\rangle = \int_{-\infty}^{\infty} x \cdot y \qquad \text{and} \qquad (x|y) = \langle x, \bar{y}\rangle$$

whenever $(x, y) \in L^p \times L^{p'}$. Suppose $1 < u < 2$ and set $w = u/(u-1)$. If $z \in L^u$, then $Y_u(z)$ is defined as the function y such that $\|y - y_n\|_w \to 0$, where $n \to \infty$ and

$$y_n(\theta) = \int_{-n}^{n} e^{-2\pi i\theta\beta} z(\beta)\, d\beta \qquad (-\infty < \theta < \infty).$$

3.4. *Remark.* Let L^0 denote the set of all step functions on $(-\infty, \infty)$ having compact support. Suppose $x \in L^0$; it is easily seen that $\Psi x \in L^+$ and $Y_u(\Psi x) = x$ whenever $1 < u < 2$.

3.5. Lemma. *Suppose* $1 < u < 2$. *If* $g \in \mathfrak{F}_\mathbf{v}$, *then*

$$[\wedge g]x = Y_u(g \cdot \Psi x) = Y_2(g \cdot \Psi x) \qquad \textit{when } x \in L^0.$$

Proof. From $x \in L^0$ it follows that $\Psi x \in L^+$ (see 3.4): therefore $g \cdot \Psi x \in L^u \cap L^2$. Thus $Y_u(g \cdot \Psi x) = Y_2(g \cdot \Psi x) = Y_2 \Psi([g\wedge]x)$; the last equality being obtained by setting $T = [\wedge g]$ in (2). The conclusion now follows from 3.4.

3.6. Lemma. *If* $T_p \in (t)_p$ *and* $q = p/(p-1)$, *then*

$$\langle T_p x^\cdot, y^\cdot\rangle = \langle x, T_q y\rangle \qquad \textit{when } (x, y) \in L^p \times L^q.$$

Proof. Set $B(x, y) = \langle T_p x^\cdot, y^\cdot\rangle$ and $B'(x, y) = \langle x, T_q y\rangle$. Both B and B' are continuous bilinear functionals on $L^p \times L^q$. Since the space L^0 is dense in both L^p and L^q (see 3.4), it will therefore suffice to show that B and B' coincide on $L^0 \times L^0$. To that effect, we will need the Parseval formula in the following two equivalent forms:

$$\langle x_1, x_2^\cdot\rangle = \langle \Psi x_1, \Psi x_2\rangle \qquad ((x_1, x_2) \in L^2 \times L^2), \tag{4}$$

$$\langle \Psi y_1, y_2\rangle = \langle y_1^\cdot, Y_2 y_2\rangle \qquad ((y_1, y_2) \in L^2 \times L^2) \tag{4'}$$

(see (**11**, Theorem 49 or 75); recall that $L^p = L^p(-\infty, \infty)$). Set $g = \mathbf{v}T$, and suppose that $(x, y) \in L^0 \times L^0$. From (4) and (2), therefore, we have:

$$\langle Tx^\cdot, y^\cdot\rangle = \langle g \cdot \Psi x^\cdot, \Psi y\rangle = \langle \Psi x^\cdot, g\Psi \cdot y^\cdot\rangle.$$

We now apply (4′) with $y_1 = x^\cdot$ and $y_2 = g \cdot \Psi y^\cdot$:

$$\langle Tx^\cdot, y^\cdot\rangle = \langle x, Y_2(g \cdot \Psi y)\rangle = \langle x, Ty\rangle;$$

the last equality comes from 3.5 and $T = [\wedge g]$.

3.7. *Remark.* The positive sesquilinear Hermitean form $\{(x, y) \to (x|y)\}$ on $L^+ \times L^+$ (see 3.3) makes L^+ into an inner-product space. From 3.6 it can easily be derived that $\sim T$ is the Hilbert adjoint of T:

$$(Tx|y) = (x|\sim Ty) \quad \text{when } x \in L^+ \text{ and } y \in L^+$$

We will make no use of these properties.

3.8. *Definition.* Suppose $-\infty < \alpha < \infty$. If $\phi \in \mathfrak{F}$, then $\tau_\alpha\phi$ will denote the function g defined for all θ in $(-\infty, \infty)$ by the relation $g(\theta) = \phi(\theta - \alpha)$.

3.9. THEOREM. *Suppose $-\infty < \alpha < \infty$. If $\phi \in \mathfrak{F}_{\mathbf{V}}$ then $\tau_\alpha\phi \in \mathfrak{F}_{\mathbf{V}}$.*

Proof. Let Ψ_α be the function $\{\theta \to e^{2\pi i\theta\alpha}\}$. Set $T^{(1)} = [\mathsf{A}\phi]$, and let T be the operator defined by the relation

$$Tx = \bar{\Psi}_\alpha \cdot T^{(1)}(\Psi_\alpha \cdot x) \qquad \text{(all } x \text{ in } L^+\text{)}.$$

Note that $|T|_p = |T^{(1)}|_p$, and therefore $T \in \mathcal{E}$. Since $g = \tau_\alpha\phi \in \mathfrak{F}$, it will suffice to show that (2) holds; but this follows easily from a repeated application of the relation $\tau_\alpha(\Psi\phi) = \Psi(\bar{\Psi}_\alpha \cdot \phi)$.

4. Two lattices of projectors. The Hilbert transformation H is defined for all x in L^+ by the relation

$$(Hx)(\theta) = \int_{-\infty}^{\infty} \frac{1}{\pi(\beta - \theta)} x(\beta)d\beta \qquad (-\infty < \theta < \infty),$$

the integral being taken in the Cauchy principal value sense. It is well known that $H \in \mathcal{E}$. The fact that $H \in t(-i\cdot\text{sgn})$ is explicitly stated in (**12**, p. 22) and (**3**, p. 8); it can be extracted from (**11**, pp. 120–125). Thus $H \in (t)$ and $H = -i\cdot\text{sgn} \in \mathfrak{F}_{\mathbf{V}}$. Since $\mathfrak{F}_{\mathbf{V}}$ is a linear space containing the function $I^0 = \{\theta \to 1\}$, it follows that $g_0 = 2^{-1}(I^0 + \text{sgn}) \in \mathfrak{F}_{\mathbf{V}}$.

Suppose that α and β belong to the closed interval $[-\infty, \infty]$. Let $I_\#^0(\alpha, \beta)$ denote the characteristic function of the open interval (α, β), and set $\phi_\alpha = I_\#^0(\alpha, \infty)$. Recall that $g_0 = 2^{-1}(I^0 + \text{sgn}) \in \mathfrak{F}_{\mathbf{V}}$, and note that $g_0 = I_\#^0(0, \infty)$. From 3.9 it can therefore be inferred that $\tau_\alpha g_0 = \phi_\alpha \in \mathfrak{F}_{\mathbf{V}}$.

4.1. *Remark.* We now know that $\mathfrak{F}_{\mathbf{V}}$ contains the function $I_\#^0(\alpha, \infty)$ whenever $\alpha \in [-\infty, \infty]$. Again using the fact that $\mathfrak{F}_{\mathbf{V}}$ is an algebra containing I^0, we deduce that $\mathfrak{F}_{\mathbf{V}}$ contains any function of the form $I_\#^0(\alpha, \beta)$, where $-\infty \leqslant \alpha \leqslant \beta \leqslant \infty$.

4.2. *Notation.* Let V denote the set of all complex-valued functions defined on $(-\infty, \infty)$ such that $|f|_v \neq \infty$, where $|f|_v$ denotes the total variation of f on $(-\infty, \infty)$. We will write

$$||f||_\infty = \sup\{|f(\theta)|: -\infty < \theta < \infty\},$$

and

$$||f||_0 = ||f||_\infty + |f|_v.$$

4.3. LEMMA. *If $L^1 \cap V$ denotes the set of all g in $L^1(-\infty, \infty)$ such that $g \in V$, then $L^1 \cap V \subset \mathfrak{F}_V$. Moreover, there exists a number $c_p > 0$ with the property that, if $g \in L^1 \cap V$, then*

$$|[\wedge g]|_p < 2^{-1}c_p|g|_v. \tag{5}$$

Proof. An operator Tg corresponds to g so that $||(Tg)x||_p < 2^{-1}c_p|g|_v||x||_p$ for all x in L^0 (see (**8**, 3.3 and 3.7), where $g = a$). Since L^0 is dense in L^+, it follows that Tg has an extension T_+ with $T_+ \in (t^+)$ and $|T_+|_p < 2^{-1}c_p|g|_v$, whence $T_+ \in \mathscr{E}$. Since $g \in \mathfrak{F}$, it remains to show that $T_+ \in t(g)$. From (**8**, 7.2 (14)) it follows that

$$\Psi(B_2(x, g)) = g \cdot \Psi x \qquad \text{(when } x \in L^0\text{)}.$$

From the definition (**8**, §5) of $B_p(x, g)$ it results immediately that $B_2(x, g) = (Tg)x$ when $x \in L^0$; consequently $B_2(x, g) = T_+x$ when $x \in L^+$. Thus $T_+ \in t(g)$, which concludes the proof.

4.4. *Remark.* Let "$<$" be the relation defined on $\mathscr{E}$ by:

$$T^{(1)} < T^{(2)} \Leftrightarrow T^{(1)} = T^{(1)}T^{(2)}.$$

A family $\mathscr{P}$ will be called an "$\mathscr{E}$-tower" if $(\mathscr{P}, <)$ forms a lattice of self-adjoint (see 2.5), idempotent members of $\mathscr{E}$ satisfying the following two conditions:

(ii) *The order-type of $(\mathscr{P}, <)$ is the order-type of some closed subinterval of $[-\infty, \infty]$;*

(iii) *If $P \in \mathscr{P}$, then $\mathbf{O} \in \mathscr{P}$ and $\mathbf{O} < P < \mathbf{I} \in \mathscr{P}$.*

4.5. Both families $\{[\wedge I_\sharp^0(\alpha, \infty)]: \alpha \in [-\infty, \infty]\}$ and $\{[\wedge I_\sharp^0(-n, n)]: 0 < n < \infty\}$ are $\mathscr{E}$-towers; in Part II it will be shown that they are the spectral resolutions pertaining to two unbounded operators.

Set $\psi_n = I_\sharp^0(-n, n)$. We here examine more closely the $\mathscr{E}$-tower $\{[\wedge\psi_n]: 0 < n < \infty\}$. Suppose $0 < n < \infty$, and let χ_n be the function defined by

$$\chi_n(\theta) = (\sin 2\pi n\theta)/\pi\theta \qquad (-\infty < \theta < \infty).$$

The Dirichlet operator $J^{(n)}$ is defined for all x in L^+ by the relation

$$(J^{(n)}x)(\theta) = \int_{-\infty}^{\infty} \chi_n(\theta - \beta)x(\beta)\, d\beta.$$

It is well known that $J^{(n)} \in \mathscr{E}$ (see (**6**)), and from (**11**, Theorem 65) we see that $\Psi(J^{(n)}x) = \Psi(\chi_n * x) = (\Psi\chi_n)\cdot(\Psi x)$. But $\Psi\chi_n = \psi_n$; therefore $J^{(n)} = [\wedge\psi_n]$.

4.6. LEMMA. *If $f \in V$ and $\psi_n = I_\sharp^0(-n, n)$, then*

$$|[\wedge f]|_p < 2^{-1}c_p \sup\{|\psi_n \cdot f|_v: 0 < n < \infty\}.$$

Proof. Clearly $h_n = \psi_n \cdot f \in L^1 \cap V$; from 4.3 therefore

$$|[\wedge h_n]|_p < 2^{-1}c_p \sup\{|\psi_n \cdot f|_v: 0 < n < \infty\} = k_p'. \tag{6}$$

Suppose $x \in L^+$, and note that

$$\lim\|[\mathsf{A}f]x - J^{(n)}([\mathsf{A}f]x)\|_p = 0 \qquad (n \to \infty) \tag{7}$$

(see, for example, **(6** (1b″)**)** or **(8**, 5.2**)**). In 4.5 we saw that $J^{(n)} = [\mathsf{A}\psi_n]$; therefore $J^{(n)} \circ [\mathsf{A}f] = [\mathsf{A}(\psi_n \cdot f)] = [\mathsf{A}h_n]$ (from (3)). Accordingly, (6) now states that $\|J^{(n)}([\mathsf{A}f]x)\|_p < k_p'\|x\|_p$, which (from (7)) gives the conclusion $\|[\mathsf{A}f]x\|_p < k_p'\|x\|_p$.

4.7. Theorem. *If $f \in V$, then $f \in \mathfrak{F}_{\mathbf{V}}$ and*

$$|[\mathsf{A}f]|_p < c_p\|f\|_0. \tag{8}$$

Proof. Suppose $0 < n < \infty$ throughout, and set $\mathfrak{a}_n = (-n, n)$, while $\mathfrak{a}_n^- = (-\infty, -n]$ and $\mathfrak{a}_n^+ = [n, \infty)$. Note first that $h_n = I_f{}^0\mathfrak{a}_n$ vanishes outside of $\mathfrak{a}_n$, so that $|h_n|_v < 2\|f\|_\infty + |f|_v$. In the notation of 4.6, we can write $h_n = \psi_n \cdot f$; consequently, the relation (8) follows from 4.6. It remains to show that $f \in \mathfrak{F}_{\mathbf{V}}$. Define $f^{(n)} = h_n + g^{(n)}$, where

$$g^{(n)} = f(-n)I_f{}^0(\mathfrak{a}_n^-) + f(n)I_f{}^0(\mathfrak{a}_n^+).$$

Since $g^{(n)}$ is a linear combination of members of $\mathfrak{F}_{\mathbf{V}}$ (see 4.1), it follows that $g^{(n)} \in \mathfrak{F}_{\mathbf{V}}$. Since $h_n \in L^1 \cap V$ and 4.3, this in turn necessitates that $f^{(n)} \in \mathfrak{F}_{\mathbf{V}}$. Set $T^{(n)} = [\mathsf{A}f^{(n)}]$ and apply (8):

$$|T^{(n)} - T^{(m)}|_p < c_p\|f^{(n)} - f^{(m)}\|_0 \qquad (m > 0). \tag{9}$$

Let $v(g; \mathfrak{a})$ denote the total variation of g on $\mathfrak{a}$; observe that $v(f - f^{(n)}; \mathfrak{a}) = v(f; \mathfrak{a})$ when $\mathfrak{a} = \mathfrak{a}_n^-$ or $\mathfrak{a} = \mathfrak{a}_n^+$. Moreover, $f - f^{(n)}$ vanishes on $\mathfrak{a}_n$, and therefore

$$\|f - f^{(n)}\|_\infty < |f - f^{(n)}|_v = v(f; \mathfrak{a}_n^-) + v(f; \mathfrak{a}_n^+).$$

Since $f \in V$, this inequality implies that

$$0 = \lim\|f - f^{(n)}\|_0 = \lim\|f - f^{(n)}\|_\infty \qquad (n \to \infty). \tag{10}$$

From (9) and (10) it can be inferred that the sequence $\{T_p{}^{(n)}\}_n$ is a Cauchy sequence in $\mathscr{E}_p$, and it accordingly converges (when $n \to \infty$) to a member T_p of $\mathscr{E}_p$. Therefore, $p \in (1, \infty)$ and $x \in L^+$ implies that $0 = \lim\|T_px - T^{(n)}x\|_p$ $(n \to \infty)$; but this in turn implies that $\{T^{(n)}x\}_n$ converges in measure to T_px. Since measure-limits are uniquely defined, the outcome can be stated as follows: $p \in (1, \infty)$ and $x \in L^+$ implies that $T_2x = T_px \in L^p$. From this we infer that $T_2 \in \mathscr{E}$ (see § 2).

The proof is now concluded by showing that $T_2 \in \mathfrak{t}(f)$. Suppose $x \in L^+$, set $\phi = \Psi(T_2x) - f \cdot \Psi x$ and note that

$$\|\phi\|_2 < |T_2 - T^{(n)}|_2\|x\|_2 + \|f - f^{(n)}\|_\infty\|x\|_2.$$

From (10) it follows that $\phi = \mathbf{O} = \Psi(T_2x) - f \cdot \Psi x$. This shows that $T_2 \in \mathfrak{t}(f)$, whence $f \in \mathfrak{F}_{\mathbf{V}}$.

4.8. Corollary. $V \subset \mathfrak{F}_{\mathbf{V}}$.

5. Two convergence theorems. Let F be a function defined on a set S. If $(S, \gg)$ is a directed set, then the net $(F, \gg)$ is also denoted $\{F(s): s \in S, \gg\}$ (our terminology and notation come from **(5**, p. 65**)**). If F maps into a set $\mathfrak{X}$, then $(F, \gg)$ is called a ***net in*** $\mathfrak{X}$. If $(F, \gg)$ is a net in a Hausdorff space $\mathfrak{X}$, then we write

$$x = \mathfrak{X} \lim\{F(s): s \in S, \gg\}$$

to indicate that $(F, \gg)$ converges to a point x in $\mathfrak{X}$ (see **(5**, p. 68**)**). Let $\mathcal{T}_p$ denote the strong operator-topology of the algebra $\mathcal{E}_p$ which was defined in § 3. For example, suppose that $F(s) \in \mathcal{E}$ (for all s in S) and $T \in \mathcal{E}$; then $F(s)$ and T admit continuous extensions $F(s)_p$ and T_p, respectively (see § 3; $F(s)_p \in \mathcal{E}_p$ and $T_p \in \mathcal{E}_p$). Accordingly, the statement

$$T_p = \mathcal{T}_p \lim\{F(s)_p: s \in S, \gg\} \tag{11}$$

means that the net $\{F(s)_p: s \in S, \gg\}$ converges to T_p in the strong operator-topology of $\mathcal{E}_p$ (see **(4**, p. 53**)**).

5.1. *Definition.* Let $(F, \gg)$ be a net in $\mathcal{E}$. If $T \in \mathcal{E}$, then

$$T = \mathcal{T} \lim\{F(s): s \in S, \gg\}$$

is written to mean that relation (11) occurs whenever $1 < p < \infty$.

5.2. *Remark.* If $\{f(s): s \in S, \gg\}$ is a net in $[0, \infty)$, then

$$\infty \neq \lim\sup\{f(s): s \in S, \gg\}$$

if and only if there exists a number $N_0 > 0$ and an element s_0 of S such that $f(s) < N_0$ whenever $s \in S$ and $s \gg s_0$.

5.3. THEOREM. *Suppose $g \in \mathfrak{F}_V$, and let $\{G(s): s \in S, \gg\}$ be a net in V. Set $\mathfrak{X}_p = L^p(-\infty, \infty)$ and suppose further that the relation*

$$[\wedge g]x = \mathfrak{X}_2 \lim\{[\wedge G(s)]x: s \in S, \gg\} \tag{12}$$

holds for all x in L^0. If

$$\infty \neq \lim\sup\{||G(s)||_0: s \in S, \gg\}, \tag{13}$$

then

$$[\wedge g] = \mathcal{T} \lim\{[\wedge G(s)]: s \in S, \gg\}.$$

Proof. Suppose $1 < p < \infty$. We must prove (11) for $T = [\wedge g]$ and $F(s) = [\wedge G(s)]$; that is, we must show that

$$T_p x = \mathfrak{X}_p \lim\{F(s)_p x: s \in S, \gg\} \tag{14}$$

for all x in $\mathfrak{X}_p$. From (13), 5.2, and 4.7 follows the existence of a number N_0 and an element s_0 of S such that, if $s \in S$ and $s \gg s_0$, then

$$|F(s)_q|_q < N_0 c_q \tag{iv}$$

whenever $1 < q < \infty$. It will be convenient to describe (iv) by saying that the net $\{F(s)_q: s \in S, \gg\}$ is e.u.b. (eventually uniformly bounded) in $\mathscr{E}_q$. Consequently, the net $\{F(s)_p: s \in S, \gg\}$ is e.u.b. in $\mathscr{E}_p$. It is easily verified that the Banach–Steinhaus theorem (**4**, p. 41) applies not only to uniformly bounded sequences in $\mathscr{E}_p$, but also to e.u.b. nets in $\mathscr{E}_p$. Let us suppose for a moment that (14) holds for all x in L^0; since L^0 is dense in $\mathfrak{X}_p$, the Banach–Steinhaus theorem implies that (14) holds for all x in $\mathfrak{X}_p$, and the theorem is proved.

Suppose $x \in L^0$, and set $y(s) = Tx - F(s)x$; in view of our preceding remark, it will suffice to show that

(v) $$0 = \lim\{||y(s)||_p: s \in S, \gg\}.$$

If $p = 2$, there is nothing to prove, since (v) is then our hypothesis (12). If $p \neq 2$ there clearly exists a number q with $1 < q < \infty$ such that p lies between 2 and q; there exists therefore a number m such that

$$\frac{1}{p} = \frac{1}{2}m + \frac{1}{q}(1 - m) \qquad \text{and} \qquad 0 < m < 1.$$

From the logarithmic convexity of the norm we see that

$$||y(s)||_p \leqslant (||y(s)||_2)^m \cdot (||Tx - F(s)x||_q)^{1-m}.$$

Accordingly, we can infer from (iv) that, if $s \gg s_0$, then

$$||y(s)||_p \leqslant (||y(s)||_2)^m \cdot ([|T|_q + N_q c_q] \,.\, ||x||_q)^{1-m}.$$

Consequently, (v) results from the hypothesis (12).

5.4. Corollary. *Suppose $g \in \mathfrak{F}_V$ and let $\{G(s): s \in S, \gg\}$ be a net in V satisfying* (13). *If*

(15) $$0 = \lim\{||g - G(s)||_\infty: s \in S, \gg\},$$

then

(16) $$[\mathsf{A}g] = \mathscr{T}\lim\{[\mathsf{A}G(s)]: s \in S, \gg\}.$$

Proof. In view of 5.3, it will suffice to establish (12). Take x in L^0; from 3.5 it follows that

$$||[\mathsf{A}g]x - [\mathsf{A}G(s)]x||_2 = ||Y_2([g - G(s)] \cdot \Psi x)||_2.$$

But $[g - G(s)] \cdot \Psi x$ is in L^+ (see 3.4). Since Y_2 is an isometric mapping, we see that

(17) $$||[\mathsf{A}g]x - [\mathsf{A}G(s)]x||_2 \leqslant ||g - G(s)||_\infty \,.\, ||\Psi x||_2.$$

The conclusion (12) now results from (15), (17), and $\infty \neq ||\Psi x||_2$.

6. The main result. From now on, $R = (-\infty, \infty)$ and $\bar{R} = [-\infty, \infty] = R \cup \{-\infty, \infty\}$; if α and β lie in $\bar{R}$, then $(\alpha, \beta] = \{\theta \in R: \alpha < \theta \leqslant \beta\}$. The space $\bar{R}^2 = \bar{R} \times \bar{R}$ consists of all points $\lambda = (\lambda_1, \lambda_2)$ such that $\lambda_1 \in \bar{R}$ and

$\lambda_2 \in \bar{R}$. The usual embedding $\{\alpha \to (\alpha, 0)\}$ of $\bar{R}$ into $\bar{R}^2$ will be assumed. Accordingly, $\bar{R} \subset \bar{R}^2$; if α and β belong to $\bar{R}^2$, then $(\alpha, \beta]$ is the Cartesian product $(\alpha_1, \beta_1] \times (\alpha_2, \beta_2]$, with the exception $(\alpha, \beta] = (\alpha_1, \beta_1] \times \{0\} = (\alpha_1, \beta_1]$ in the case $\alpha = \alpha_1$ and $\beta = \beta_1$.

6.1. *Definitions.* If $Q \subset \bar{R}^2$, then $\mathfrak{B}Q$ will denote the family of all finite unions of members of $\mathfrak{A}Q = \{(\alpha, \beta]: (\alpha, \beta) \in Q \times Q\}$.

6.2. The Boolean algebra $\mathfrak{C}_\blacktriangle$ will consist of all symmetric differences $B \dotplus N = (B \cup N) - (B \cap N)$, where $B \in \mathfrak{B}\bar{R}$ and N is a subset of R having zero measure.

6.3. The following notations will be used consistently. If $g \in \mathfrak{F}$, then $g_1 =$ (real part of g) and $g_2 =$ (imaginary part of g). If $\sigma \in \mathfrak{B}\bar{R}^2$, then $(g \in \sigma) = \{\theta \in R: g(\theta) \in \sigma\}$, except that $(g \in \sigma) = (g_1 \in \sigma)$ whenever $g = g_1$.

6.4. The set $\mathfrak{F}_\blacktriangle$ will consist of all functions g in $\mathfrak{F}$ such that $(g \in \sigma) \in \mathfrak{C}_\blacktriangle$ whenever $\sigma \in \mathfrak{A}\bar{R}^2$.

6.5. If $T \in (t)$ and $g = \mathrm{v}T \in \mathfrak{F}_\blacktriangle$, then the set-function E^T is defined for all σ in $\mathfrak{B}\bar{R}^2$ by the relation

$$E^T(\sigma) = [\mathbf{A}I_\sharp^0(g \in \sigma)].$$

Recall that $\psi = I_\sharp^0(g \in \sigma)$ is a function such that $\psi(\theta) = 1$ whenever $\theta \in (g \in \sigma)$, while $\psi(\theta) = 0$ otherwise. Note that $\psi \in V$; in this connection, it should also be mentioned that $\mathfrak{A}\bar{R}$, $\mathfrak{B}\bar{R}$, and $\mathfrak{C}_\blacktriangle$ are Boolean set-algebras. Since the verification of these facts is routine, it will be omitted. Both $\emptyset$ and R^2 belong to $\mathfrak{B}\bar{R}^2$; it is clear that

$$E^T(\emptyset) = \mathbf{O} \quad \text{and} \quad E^T(R^2 - \sigma) = \mathbf{I} - E^T(\sigma)$$

whenever $\sigma \in \mathfrak{B}\bar{R}^2$. In fact, E^T is an isomorphism into (t) of the Boolean set-algebra $\mathfrak{B}\bar{R}^2$; if σ' and σ'' are in $\mathfrak{B}\bar{R}^2$, then

$$E^T(\sigma' \cup \sigma'') = E^T(\sigma') \vee E^T(\sigma'')$$

and

$$E^T(\sigma' \cap \sigma'') = E^T(\sigma') \wedge E^T(\sigma'')$$

(the operations "$\vee$" and "$\wedge$" are defined in (**1**, p. 219)).

6.6. *Orientation.* The following is aimed at defining two-dimensional Stieltjes integrals of commonplace type. In order to implement a later proof (6.14), an order-preserving notation for range partitions will first be described.

6.7. Let $\mathfrak{Z}$ be the family of all strictly monotone-increasing functions Z whose domain $D(Z)$ is a finite set of consecutive integers, and whose range $\{Z_\nu: \nu \in D(Z)\}$ is a subset of $\bar{R}$. If $Z \in \mathfrak{Z}$, we denote by Z^* the set $\{\nu \in D(Z):$

$\nu > \min D(Z)\}$ and write $Z(\nu] = (Z_{\nu-1}, Z_\nu]$ whenever $\nu \in Z^*$. In case $Q_\iota \subset \bar{R}$, then $\mathfrak{Z}Q_\iota$ will denote the family of all Z in $\mathfrak{Z}$ such that

$$Q_\iota \subset \bigcup \{Z(\nu] : \nu \in Z^*\}.$$

6.8. *Definition.* Suppose $g \in \mathfrak{F}$, and denote by $[g]$ the closed cell $[-\lambda, \lambda]$, where $\lambda_\iota = \|g_\iota\|_\infty$ for $\iota = 1, 2$ (see 4.2). The family $S[g]$ consists of all ordered pairs $(Z, \mathfrak{z})$ whose first member $Z = (Z_1, Z_2)$ lies in $\mathfrak{Z}[g_1] \times \mathfrak{Z}[g_2]$, and such that $\mathfrak{z}$ is a function on $Z^* = Z_1^* \times Z_2^*$ whose values $\mathfrak{z}(\nu)$ lie in $Z^\nu = Z_1(\nu_1] \times Z_2(\nu_2]$ whenever $\nu = (\nu_1, \nu_2) \in Z^*$.

6.9. *Definition.* Suppose $T \in (t)$ and $\mathbf{v}T \in \mathfrak{F}_\blacktriangle$. If $s = (Z, \mathfrak{z}) \in S[\mathbf{v}T]$, then we write

$$(E^T : s) = \sum_{\nu \in \omega} \mathfrak{z}(\nu) E^T(Z^\nu) \qquad (\omega = Z^*).$$

6.10. THEOREM. *Suppose $T \in (t)$ and $\mathbf{v}T \in \mathfrak{F}_\blacktriangle$. If there exists a number $k_0 > 0$ such that $|\mathbf{v}(E^T : s)|_p < k_0\|\mathbf{v}(E^T : s)\|_\infty$ whenever $s \in S[\mathbf{v}T]$, then the following Stieltjes integral exists:*

$$\int \lambda \cdot E^T(d\lambda) = \mathscr{T} \lim \{(E^T : s) : s \in S[\mathbf{v}T], \gg\}. \tag{18}$$

Moreover,

$$T = \int \lambda \cdot E^T(d\lambda). \tag{1}$$

6.11. *Remarks.* The set $S[\mathbf{v}T]$ is directed by the partial ordering "$\gg$" (see (**5**, p. 79) and 6.12). The meaning of the relation (1) will now be explicitly formulated. If $1 < p < \infty$, then the net

$$\left\{\left\| T_p x - \sum_{\nu \in \omega} \mathfrak{z}(\nu) E^T(Z^\nu)_p x \right\|_p : (Z, \mathfrak{z}) \in S[\mathbf{v}T], \gg\right\}$$

converges to zero for all x in $L^p(R)$ (compare (18) with 5.1). Consequently, (1) implies that the net

$$\left\{\sum_{\nu \in \omega} \mathfrak{z}(\nu) E^T(Z^\nu)_p : (Z, \mathfrak{z}) \in S[\mathbf{v}T], \gg\right\}$$

converges to T_p in the weak operator-topology (this again comes from (18) and 5.1); T_p is therefore a "scaled" member of $\mathscr{E}_p$ (see (**7**, p. 450)).

6.12. *Proof of* 6.10. If $Q \subset R^2$, let $|Q|$ denote the diameter of Q. Set $g = \mathbf{v}T$ and $S = S[g]$. Suppose $s = (Z, \mathfrak{z}) \in S$. We define $\|s\| = \max\{|Z^\nu| : \nu \in Z^*\}$. The partial ordering is defined by: $s' \gg s \Leftrightarrow \|s'\| < \|s\|$ whenever $s' \in S$. Set $G(s) = \mathbf{v}(E^T : s)$; from 6.9 and 6.5 we note that

$$G(s) = \sum_{\nu \in \omega} \mathfrak{z}(\nu) I_\beta^0(g \in Z^\nu) \qquad (\omega = Z^*). \tag{19}$$

Clearly $G(s) \in V$ (see 6.5). It is easily seen that

$$\|g - G(s)\|_\infty < \|s\|. \tag{20}$$

But $\infty \neq ||g||_\infty$ and therefore $\infty \neq \lim\sup\{||G(s)||_\infty : s \in S, \gg\}$ (see 5.2), from which our hypothesis $||G(s)||_0 < (k_0 + 1)||G(s)||_\infty$ yields the relation (13) of 5.3. Since (20) implies (15) in 5.4, the net $\{G(s): s \in S, \gg\}$ satisfies all the conditions of 5.4. The conclusion now results from (16), $T = [\wedge g]$ and $(E^T: s) = [\wedge G(s)]$.

6.13. *Definition.* A function f is "piecewise monotone" if there exists a member Z of $\mathfrak{Z}\bar{R}$ such that f is monotone on $Z(\nu]$ for all ν in Z^* (see 6.7).

6.14. THEOREM. *Let g be a bounded function whose real and imaginary parts are piecewise monotone. Then $g \in \mathfrak{F}_\blacktriangle$ and $[\wedge g]$ is a member T of (t) such that*

$$T = \int \lambda \cdot E^T(d\lambda) \tag{1}$$

in the sense of 6.10–6.11.

COROLLARY. *Suppose that A is a bounded Radon measure on R, and let g be the Fourier transform of A. If T_p is the convolution operator $A *_p$, then T satisfies* (1) *whenever g satisfies the hypothesis of* 6.14.

Proof. Observe that $g = \Psi(dA)$ in the notation of 3.1; from 3.2 therefore $\mathbf{v}T = g$, and the conclusion now comes from 6.14.

6.15. *Remark.* Suppose $J \in \mathfrak{A}\bar{R}$, and let f belong to the set $\mathfrak{G}(J)$ of all real-valued functions that are monotone increasing on J. If $\sigma = (\alpha, \infty)$ or $\sigma = [\alpha, \infty)$, then $J \cap (f \in \sigma)$ is a connected subset of $\bar{R}$; therefore $J \cap (f \in \sigma) \in \mathfrak{C}_\blacktriangle$.

6.16. Consider now the case $\sigma = (\alpha, \beta] \in \mathfrak{A}\bar{R}$; then $J \cap (f \in \sigma) \in \mathfrak{C}_\blacktriangle$. This can be seen by noting that $(f \in \sigma)$ is the set-theoretic difference $J \cap (f \in \sigma_1) - J \cap (f \in \sigma_2)$, where $\sigma_1 = (\alpha, \infty)$ and $\sigma_2 = (\beta, \infty)$; since $\mathfrak{C}_\blacktriangle$ is a Boolean ring, the conclusion follows from 6.15.

6.17. *Definition.* If $J \in \mathfrak{A}\bar{R}$, then $\mathfrak{M}(J)$ will be the set of all bounded functions whose real and imaginary parts are both monotone on J.

6.18. LEMMA. *If $J \in \mathfrak{A}\bar{R}$ and $g \in \mathfrak{M}(J)$, then $J \cap (g \in \sigma) \in \mathfrak{C}_\blacktriangle$ whenever $\sigma \in \mathfrak{A}\bar{R}^2$.*

Proof. Since $\sigma \in \mathfrak{A}\bar{R}^2$, we can write $\sigma = \sigma_1 \times \sigma_2$, where $\{\sigma_1, \sigma_2\} \subset \mathfrak{A}\bar{R}$, so that $J \cap (g \in \sigma) = J \cap (g_1 \in \sigma_1) \cap (g_2 \in \sigma_2)$. Set $\iota = 1, 2$. The proof will therefore be concluded by establishing that $J \cap (g_\iota \in \sigma_\iota) \in \mathfrak{C}_\blacktriangle$. Since this was proved in 6.16 for the case $g_\iota \in \mathfrak{G}(J)$, it will suffice to consider the case where g_ι is decreasing on J. But then $f = -g_\iota \in \mathfrak{G}(J)$, and the arguments in 6.16 (together with 6.15), give the conclusion $J \cap (g_\iota \in \sigma_\iota) \in \mathfrak{C}_\blacktriangle$.

6.19. *Definition.* If $Q \subset R^2$, then $\mathfrak{U}Q$ will denote the set of all mappings F of Q into R^2 such that, if $\lambda' = (\lambda_1', \lambda_2') \in Q$ and $\lambda'' = (\lambda_1'', \lambda_2'') \in Q$, then $\lambda_\iota' < \lambda_\iota''$ implies $F_\iota(\lambda') < F_\iota(\lambda'')$ whenever $\iota = 1$ and also when $\iota = 2$.

6.20. LEMMA. *Suppose $J \in \mathfrak{A}\bar{R}$ and $g \in \mathfrak{M}(J)$. If $F \in \mathfrak{U}[g]$ then $(F \circ g) \in \mathfrak{M}(J)$.*

Proof. The composition $(F \circ g)$ is the function h such that $h(\theta) = F(g(\theta))$ for all θ in R. In case $\theta' < \theta''$ and $g_1(\theta') < g_1(\theta'')$, set $\lambda' = g(\theta')$ and $\lambda'' = g(\theta'')$; then $\lambda_1' < \lambda_1''$ and $F_1(g(\theta')) < F_1(g(\theta''))$. Therefore $h_1 \in \mathfrak{G}(J)$. The remaining cases can be similarly derived.

6.21. *Remark.* Let $h \in \mathfrak{F}$ and $J = (\alpha, \beta] \in \mathfrak{A}\bar{R}$. Denote by $v(h; J)$ the total variation of h on $[\alpha, \beta] \cap R$. If $h \in \mathfrak{M}(J)$ (see 6.17), it is easily verified that $v(h; J) < 8\|h\|_\infty$.

Proof of 6.14. Set $\iota = 1, 2$. By hypothesis there exist two members Π_1 and Π_2 of $\mathfrak{Z}\bar{R}$ (see 6.7) such that g_ι is monotone on each $\Pi_\iota(\kappa_\iota]$ when $\kappa_\iota \in \Pi_\iota^*$. For any $\kappa = (\kappa_1, \kappa_2)$ in $\Pi^* = \Pi_1^* \times \Pi_2^*$, we write $\Pi^\kappa = \Pi_1(\kappa_1] \cap \Pi_2(\kappa_2]$. Note that $\Pi^\kappa \in \mathfrak{A}\bar{R}$ and $g \in \mathfrak{M}(\Pi^\kappa)$.

Observe first that $g \in V$, and therefore $g \in \mathfrak{F}_\blacktriangledown$ (by 4.8). Thus $\blacktriangle g = T \in (t)$ and $\blacktriangledown T = g$. The property $g \in \mathfrak{F}_\blacktriangle$ is proved as follows. Take any σ in $\mathfrak{A}\bar{R}^2$, and note that $(g \in \sigma) = \bigcup\{\Pi^\kappa \cap (g \in \sigma) : \kappa \in \Pi^*\}$; since $\mathfrak{C}_\blacktriangle$ is a Boolean ring, the conclusion $(g \in \sigma) \in \mathfrak{C}_\blacktriangle$ is now inferred from 6.18.

Next, take any $s = (Z, \mathfrak{z})$ in $S[g]$, set $G(s) = \blacktriangledown(E^T: s)$ and note that

$$|G(s)|_p < \sum_{\kappa=1}^{m} v(G(s); \Pi^\kappa), \tag{21}$$

where $\{1, 2, 3, \ldots, m\} = \Pi^*$. From Definition 6.8, there exist functions Z_ι in $\mathfrak{Z}$ such that $\mathfrak{z}(\nu) \in Z_\nu = Z_1(\nu_1] \times Z_2(\nu_2]$ for all $\nu = (\nu_1, \nu_2)$ in $Z_1^* \times Z_2^*$ (the index-sets Z^* are defined in 6.7). If $\lambda \in [g]$, denote by $\nu[\lambda]$ the ν in Z^* such that $\lambda \in Z^\nu$, and let F be the function defined by $F(\lambda) = \mathfrak{z}(\nu[\lambda])$ for all λ in $[g]$. From the isotonicity of the correspondences set up in 6.7 it now follows that $F \in \mathfrak{U}[g]$ (see 6.19). On the other hand, it is easily checked that $G(s) = (F \circ g)$ (see (19)). From 6.20 therefore: $G(s) \in \mathfrak{M}(J)$ whenever $J \in \mathfrak{A}\bar{R}$.

Suppose $\kappa \in \Pi^*$. Since $G(s) \in \mathfrak{M}(\Pi^\kappa)$, it results from 6.21 that $v(G(s); \Pi^\kappa) < 8\|G(s)\|_\infty$, and from (21) therefore: $|G(s)|_p < 8m\|G(s)\|_\infty$. In view of 6.10, the proof of 6.14 is completed.

7.0. Added in proof. Part II of this article has appeared in the Journal of Math. and Mechanics, Vol. 10 (1961), 111–134.

7.1. *Remark.* (added March 9, 1961). The set V (defined in 4.2) is strictly included in the set V_β of all functions having generalized higher β-variation; it can be proved that $V_\beta \subset \mathfrak{F}_\blacktriangledown$. This last assertion is clearly stronger than our Corollary 4.8; it is implicit in a remark on p. 242 of an article by I. I. Hirschman, Jr. "*On multiplier transformations*", Duke Math. J., 26 (1959), 221–242. At the time the present article was written, I was unaware of Professor Hirschman's remark.

References

1. N. Dunford, *A survey of the theory of spectral operators*, Bull. Amer. Math. Soc., *64* (1958), 217–274.
2. ——— *Spectral theory in abstract spaces and Banach algebras*, Proceedings of the symposium on spectral theory and differential problems (Stillwater, Oklahoma, 1955), 1–65.
3. E. Hille, *On the generation of semi-groups and the theory of conjugate functions*, Kungl. Fysiogr. Sällsk. Lund Förhandlinger, *21* (1952), 1–13.
4. E. Hille and R. S. Phillips, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publ., XXXI (1957).
5. J. L. Kelley, *General topology* (D. Van Nostrand Co., Inc., New York, 1955).
6. H. Kober, *On Dirichlet's singular integral*, Quart. J. Math. Oxford, *11* (1940), 66–80.
7. G. L. Krabbe, *Convolution operators that satisfy the spectral theorem*, Math. Zeitschr., *70* (1959), 446–462.
8. ——— *A space of multipliers of type $L^p(-\infty, \infty)$*, Pac. J. Math., *9* (1959), 729–737.
9. ——— *Spectral invariance of convolution operators on L^p $(-\infty, \infty)$*, Duke Math. J., *25* (1958), 131–141.
10. M. A. Neumark, *Involutive Algebren*, Sowjetische Arbeiten zur Funktion-analysis (Berlin 1954), 91–191; translated from the article *Kol'ca c involyuciei*, Usp. Mat. Nauk, *3* (1948), 52–145.
11. E. C. Titchmarsh, *Introduction to the theory of Fourier integrals* (Oxford University Press, 1948).
12. O. A. Varsavsky, *Sobre la transformacion de Hilbert*, Revista Unión Mat. Argentina, *14* (1949), 20–37.

Purdue University

HOMOGENEOUS CONTINUA WHICH ARE ALMOST CHAINABLE[1]

C. E. BURGESS

The only known examples of nondegenerate homogeneous plane continua are the simple closed curve, the circle of pseudo-arcs **(6)**, and the pseudo-arc **(1; 13)**. Another example, called the pseudo-circle, has been suggested by Bing **(2)**, but it has not been proved to be homogeneous. (Definitions of some of these terms and a history of results on homogeneous plane continua can be found in **(6)**.) Of the three known examples, the pseudo-arc is both linearly chainable and circularly chainable, and the simple closed curve and the circle of pseudo-arcs are circularly chainable but not linearly chainable. It is not known whether every homogeneous plane continuum is either linearly chainable or circularly chainable. Bing has shown that a homogeneous continuum is a pseudo-arc provided it is linearly chainable **(4)**.

In this paper, a study is made of continua that are almost chainable, and the effect upon them by a homogeneity requirement is considered. It is hoped that these results might be of some help in a search for other examples of homogeneous plane continua or in an attempt to characterize such continua.

Bing has shown that a homogeneous plane continuum is a simple closed curve if it contains an arc **(5)**. Some of the theorems presented here give conditions under which a nondegenerate homogeneous plane continuum would contain a pseudo-arc. Perhaps this is a property of all such continua that do not contain an arc. Continua which are almost chainable and for which each point is an end point are characterized as continua for which every nondegenerate proper subcontinuum is a pseudo-arc. It is not known whether every such continuum is homogeneous. A more general question has been raised in **(8)**.

Throughout this paper, a *continuum* denotes a compact connected metric space. Where there is no reference to a space in which a continuum under discussion is imbedded, the continuum itself is considered as space. Where a plane continuum M is being discussed, M should be considered imbedded in a plane E and some of the coverings of M might be collections of open sets in E.

Definitions. Linear chains, circular chains, trees, and continua described with them are defined in **(10)**. Various types of homogeneity are defined in **(9)**.

Received March 1, 1960. Presented to the American Mathematical Society, January 29, 1960.

[1]This work was supported in part by the National Science Foundation under G5880.

A continuum M is ***almost chainable*** if, for every positive number ϵ, there exist an ϵ-covering G of M and a linear chain C $(L_1, L_2, \ldots, L_n)$ of elements of G such that no L_i $(1 < i < n)$ intersects an element of $G - C$ and every point of M is within a distance ϵ of some element of C. The set L_1 is called an ***end link*** of G. A point p is called an ***end point*** of M if, for every positive number ϵ, there is an ϵ-covering G of M such that p is in an end link of G.

Definitions of end links of trees,[2] branches of trees, and k-branched continua are given in **(15)**. A ***junction link*** of a tree T is an element of T that intersects at least three other elements of T. A tree-like continuum is said to be ***k-junctioned***, or to have k junctions, if k is the least integer such that, for every positive number ϵ, M can be covered by an ϵ-tree with k junction links.

THEOREM 1. *If the continuum M is nearly homogeneous and almost chainable, then M has a dense set of end points.*

Proof. That M has an end point can be shown by a method similar to the proof given by Bing **(4)** to show that a continuum has an end point if it is homogeneous and linearly chainable. Then Theorem 1 follows from the near-homogeneity of M and the fact that, under a homeomorphism of M onto itself, each end point of M goes into an end point of M.

THEOREM 2. *If the continuum M is almost chainable and K is a proper subcontinuum of M which contains an end point p of M, then K is linearly chainable with p as an end point.*

Proof. Let q be a point of $M - K$, and let ϵ be a positive number that is less than the distance from q to K. There exist an $\epsilon/2$-covering G of M and a linear chain $C(L_1, L_2, \ldots, L_n)$ in G such that: (1) no $L_i (1 < i < n)$ intersects an element of $G - C$; (2) every point of M is within a distance $\epsilon/2$ of some element of C; and (3) p is in L_1. There exists an element L_j of C such that the distance from q to L_j is less than $\epsilon/2$, and it follows that K does not intersect L_j. This implies that K is covered by the linear ϵ-chain $(L_1, L_2, \ldots, L_{j-1})$. Thus K is linearly chainable with p as an end point.

THEOREM 3. *In order that every nondegenerate proper subcontinuum of the continuum M should be a pseudo-arc, it is necessary and sufficient that M be almost chainable with each of its points as an end point.*

Proof of sufficiency. Let K be a nondegenerate proper subcontinuum of M and let p be a point of K. By Theorem 2, K is linearly chainable with p as an end point. It follows from Theorem 16 of **(3)** that K is a pseudo-arc.

The following lemma will be used in proving that the condition is necessary.

LEMMA 3.1. *If every nondegenerate proper subcontinuum of the continuum M is a pseudo-arc, K is a pseudo-arc in M, $C(L_1, L_2, \ldots, L_n)$ is a linear chain*

[2]A collection that is called a tree in **(10)** is called a tree-chain in **(15)**.

which is an essential covering of K, and p is a point of $K - (L_1 + L_n)$, then there is a linear chain $C'(L_1', L_2', \ldots, L_n')$ such that: (1) *for each $i(1 \leq i \leq n)$, L_i' is a subset of L_i*; (2) *for each $i(1 < i < n)$, the boundary of L_i' does not contain a point of M that is not covered by C'; and* (3) *p is in an element of C'.*[3]

Proof of Lemma 3.1. Let K' be a component of $K - (L_1 + L_n)$ that intersects both $\mathrm{cl}(L_1)$ and $\mathrm{cl}(L_n)$, and let K'' be the component of $K - (L_1 + L_n)$ that contains p. Let A denote the closed set $M - (L_1 + L_n)$ and let B denote the closed set $M - (L_1 + L_2 + \ldots + L_n)$. Now suppose that some continuum H in A intersects both $K' + K''$ and B. This leads to the contradiction that $H + K$ is decomposable. Hence it follows from **(14**, Theorem 35, p. 21**)** that A is the sum of two mutually separated closed sets A_1 and A_2 containing $K' + K''$ and B, respectively. Let L_1' and L_n' denote L_1 and L_n, respectively, and for each $i(1 < i < n)$, let $L_i' = A_1 \cdot L_i$. The chain $C'(L_1', L_2', \ldots, L_n')$ satisfies the conclusion of Lemma 3.1.

Proof of necessity. Since every proper subcontinuum of M is indecomposable **(1**; **12)**, it follows that M is indecomposable. Let ϵ be a positive number and let p be a point of M. There exists a pseudo-arc K in M such that p is in K and every point of M is within a distance $\epsilon/2$ of K. Let $D(R_1, R_2, \ldots, R_t)$ be a linear $\epsilon/2$ chain which is an essential covering of K such that p is in R_1. It follows from the proof of Theorem 13 of **(1)** that there exists a linear chain $C(L_1, L_2, \ldots, L_n)$ which is a refinement of D such that: (1) C is an essential covering of K; and (2) L_1 and L_n are subsets of R_t. By Lemma 3.1, there exists a linear chain $C'(L_1', L_2', \ldots, L_n')$ such that: (1) for each $i(1 \leq i \leq n)$, L_i' is a subset of L_i; (2) for each $i(1 < i < n)$, the boundary of L_i' does not contain a point of M that is not covered by C'; and (3) p is in an element of C'. Now for each $i(1 \leq i \leq t)$, let R_i' denote the sum of the elements of C' that lie in R_i. Let D' denote the linear chain $(R_1', R_2', \ldots, R_t')$. There exists an ϵ-covering G of M such that: (1) each link of D' is an element of G; (2) each point of M is within a distance ϵ of some link of D'; (3) no element of $G - D'$ intersects a link of D' different from R_t'; and (4) p is in R_1'. Hence M is almost chainable and each point of M is an end point of M.

THEOREM 4. *If the continuum M is circularly chainable and hereditarily indecomposable, then M is almost chainable and each point of M is an end point of M.*

Proof. Since every proper subcontinuum of M is linearly chainable and hereditarily indecomposable, it follows that every nondegenerate proper subcontinuum of M is a pseudo-arc **(2)**. Thus the conclusion of Theorem 4 follows from Theorem 3.

THEOREM 5. *If the continuum M is homogeneous and almost chainable, then every nondegenerate proper subcontinuum of M is a pseudo-arc.*

---

[3]This lemma and its proof might be compared with Property 17 and its proof in **(5)**.

Proof. Using the homogeneity of M, it can be shown by a method similar to the proof of Theorem 1 that every point of M is an end point of M. Hence it follows from Theorem 3 that every nondegenerate proper subcontinuum of M is a pseudo-arc.

THEOREM 6. *If the continuum M is almost chainable, then M is not a triod.*

Proof. Suppose that M is a triod. Let K be a subcontinuum of M such that $M-K$ is the sum of three mutually separated sets K_1, K_2, and K_3. For each i ($i \leqslant 3$), let D_i be an open set such that $\mathrm{cl}(D_i)$ is a subset of K_i. Let ϵ be a positive number such that, for each i, ϵ is less than the distance from $\mathrm{cl}(D_i)$ to $M-K_i$ and less than the distance from some point of D_i to the boundary of D_i. There exist an ϵ-covering G of M and a linear chain $C(L_1, L_2, \ldots, L_n)$ in G such that: (1) no $L_j(1 \leqslant j < n)$ intersects an element of $G-C$; and (2) every point of M is within a distance ϵ of some link of C. Hence for each $i(i \leqslant 3)$, some link L_{r_i} of C contains a point p_i of D_i and does not intersect $M-K_i$. Consider the case in which $r_1 < r_2 < r_3$. Then each of the links L_{r_1} and L_{r_3} of the linear chain C intersects the continuum $K+K_1+K_3$, but L_{r_2} does not intersect this continuum. It follows that for some integer j less than n, the continuum $K+K_1+K_3$ contains a point of the boundary of L_j that is not in a link of C. This involves the contradiction that L_j intersects an element of $G-C$. Hence M is not a triod.

Remark. While there does not exist a triod in a continuum that is linearly chainable **(10)**, there does exist a continuum which contains a triod and is almost chainable. A continuum which is the sum of a simple triod T and a ray spiralling around T is such an example.

THEOREM 7. *If the continuum M is almost chainable, then M is unicoherent.*

Proof. Suppose that M is the sum of two continua M_1 and M_2 and that p and q are two points of $M_1 \cdot M_2$. Consider the case in which, for every positive number ϵ, there exists an ϵ-covering G of M and a linear chain $C(L_1, L_2, \ldots, L_n)$ in G such that: (1) no $L_i(1 \leqslant i < n)$ intersects an element of $G-C$; (2) every point of M is within a distance ϵ of some link of C; and (3) L_1 intersects M_1. For a choice of ϵ that is sufficiently small, M_1 would be covered by C, and p and q would lie in two links L_i and L_j, respectively, of C. Hence every link of C between L_i and L_j would intersect both M_1 and M_2. That p and q lie in the same component of $M_1 \cdot M_2$, and hence that M is unicoherent, can be shown by a proof similar to the one given for Theorem 1 of **(6)**.

Remark. While a continuum is hereditarily unicoherent if it is linearly chainable **(6)**, this is not the case for continua that are almost chainable. A continuum which is the sum of a circle K and a ray spiralling around K is almost chainable but fails to be hereditarily unicoherent.

THEOREM 8. *If the continuum M is almost chainable, then M is irreducible between some two points.*

Proof. By Theorems 6 and 7, M is unicoherent and is not a triod. Sorgenfrey **(16)** has shown that such a continuum is irreducible between some two points.

Remark. Theorem 8 is a generalization of Rosen's result that a continuum is irreducible between some two points if it is linearly chainable **(15)**.

THEOREM 9. *If the continuum M is nearly homogeneous and almost chainable, then M is indecomposable.*

Proof. By Theorem 8, M is irreducible between some two points, and such a continuum is indecomposable if it is nearly homogeneous **(7)**.

THEOREM 10. *If M is an indecomposable plane continuum and, for each positive number ϵ, there exists a circular ϵ-chain of open disks covering M, then M is almost chainable.*

The following definition and lemma will be used in the proof of this theorem.

Definition. A circular chain $C(L_1, L_2, \ldots, L_m)$ is said to *fold back one revolution* in a circular chain $D(K_1, K_2, \ldots, K_n)$ if C is a refinement of D and there exist two links K_i and K_j of D and three links L_r, L_s, and L_t of C such that: (1) K_i intersects K_j; (2) L_s is a subset of K_i; (3) L_r and L_t are subsets of K_j; and (4) there is a linear chain in C that contains L_s, has L_r and L_t as end links, and has no link that intersects both K_i and K_j.

LEMMA 10.1. *If for each positive number ϵ, the continuum M can be covered by two circular ϵ-chains $C(L_1, L_2, \ldots, L_m)$ and $D(K_1, K_2, \ldots, K_n)$ such that C folds back one revolution in D, then M is almost chainable.*

Proof of Lemma 10.1. Let K_i and K_j be links of D and Let L_r, L_s, and L_t be links of C such that the requirements of the definition above are satisfied. For convenience, suppose that $i = 1$ and $j = n$. Let C' denote the linear chain in C that contains L_s, has L_r and L_t as end links, and has no link that intersects both K_1 and K_n. For each $q(1 < q < n)$, let H_q denote the sum of the elements of C' that lie in K_q. Let G denote the collection consisting of the sets $H_1, H_2, \ldots, H_n$ and the elements of $C - C'$. The collection G is an ϵ-covering of M such that: (1) no $H_q(1 < q < n)$ intersects an element of $G - C'$; and (2) each point of M is within a distance ϵ of one of the sets $H_1, H_2, \ldots, H_n$. Hence M is almost chainable.

Proof of Theorem 10. Let ϵ be a positive number. There exists a circular ϵ-chain $D(K_1, K_2, \ldots, K_n)$ of open disks covering M such that: (1) for each i $(1 < i < n)$, $\mathrm{cl}(K_i) \cdot \mathrm{cl}(K_{i+1, \bmod n})$ is a closed disk; and (2) the sum of the elements of D is an open annular ring. Let H and J be the two simple closed curves on the boundary of this annular ring. It follows from the indecomposability of M that there exist two disjoint subcontinua M_1 and M_2 of M and two consecutive links, say K_1 and K_n, of D such that M_1 and M_2 are covered by the linear chain $(K_1, K_2, \ldots, K_{n-1})$ and are irreducible from $K_1 \cdot \mathrm{cl}(K_n)$ to

$K_{n-1} \cdot \mathrm{cl}(K_n)$. Let δ be a positive number that is less than the distance from M_1 to M_2, and let $C(L_1, L_2, \ldots, L_m)$ be a circular δ-chain of open disks covering M such that each $\mathrm{cl}(L_i)$ is a subset of an element of D and the links of C satisfy conditions similar to those required for D in (1) and (2) above. Let J_n denote the boundary of K_n. Then $J_n \cdot \mathrm{cl}(K_1)$ and $J_n \cdot \mathrm{cl}(K_{n-1})$ are arcs ab and cd, respectively, where $a + c$ and $b + d$ are subsets of H and J, respectively. Let W denote the collection of all linear chains in C that are refinements of the linear chain $(K_1, K_2, \ldots, K_{n-1})$ and are irreducible[4] from ab to cd. Let $C_1, C_2, \ldots, C_r$ denote the chains of W, and for each i $(1 \leqslant i \leqslant r)$, let L_{p_i} and L_{q_i} be the end links of C_i that intersect ab and cd, respectively. It follows from the choice of δ that $r > 1$. For convenience, suppose that $p_1 = 1$ and that the chain C_1 consists of the elements $L_1, L_2, \ldots, L_{q_1}$ of C. There are two cases to consider.

Case 1. There exist two integers i and j $(1 \leqslant i < j \leqslant r)$ such that either no p_u is between q_i and q_j or no q_u is between p_i and p_j. This implies that C folds back one revolution in D, and hence it follows from Lemma 10.1 that M is almost chainable.

Case 2. The requirements of Case 1 are not satisfied. It will be shown that this case is impossible. For convenience, suppose that the sets $L_{p_1}, L_{p_2}, \ldots, L_{p_r}$ intersect the arc ab in the order named from ab. It follows from (**14**, Theorem 17, p. 167) that the sets $L_{q_1}, L_{q_2}, \ldots, L_{q_r}$ intersect the arc cd in the order named from c to d. It follows from (**14**, Theorem 17, p. 189) that there exist two disjoint arcs ef and gh that are irreducible from H to J such that: (1) $e + g$ and $f + h$ are subsets of H and J, respectively; (2) ef and gh do not intersect $\mathrm{cl}(K_n)$; (3) for each i $(1 \leqslant i \leqslant r)$, each of the arcs ef and gh intersects the closure of one and only one link of C_i and this intersection is a connected set; and (4) neither ef nor gh intersects the closure of a link of C unless that link is in one of the chains $C_1, C_2, \ldots, C_r$. Let Y be the simple closed curve formed by the arcs ef and gh and two arcs eg and gh of H and J, respectively, that do not intersect $\mathrm{cl}(K_n)$. By considering the order on Y of the intersections of the arcs ef and gh with links of the chains $C_1, C_2, \ldots, C_r$, it follows from (**14**, Theorem 17, p. 167) that if the links of $C(L_1, L_2, \ldots, L_m)$ are followed in their natural order in C, then the end links of the chains $C_1, C_2, \ldots, C_r$ would occur as follows. First, $L_{p_1} = L_1$ would occur, next L_{q_1} would occur, then some L_{p_i} $(i > 1)$ would occur, then L_{q_i} would occur, then some L_{p_j} $(j > i)$ would occur, etc. By continuing this way until L_{q_r} occurs, then some L_{p_s} $(s < r)$ would occur next, and this would involve a contradiction to (**14**, Theorem 17, p. 167).

Remark. It would be interesting to know whether every plane continuum M that is circularly chainable can be imbedded in the plane so that, for

[4]A linear chain C is *irreducible between two sets* X and Y if one end link of C intersects X and the other intersects Y but no proper subchain of C has this property.

every positive number ϵ, M can be covered by a circular ϵ-chain of open disks.[6] Every continuum M that is linearly chainable can be imbedded in the plane so that, for every positive number ϵ, M can be covered by a linear ϵ-chain of open disks **(3)**. However, there do exist continua, for example solenoids **(5)**, which are circularly chainable and cannot be imbedded in the plane.

THEOREM 11. *If M is a homogeneous indecomposable plane continuum such that, for each positive number ϵ, M can be covered by a circular ϵ-chain of open disks, then every nondegenerate proper subcontinuum of M is a pseudo-arc.*

Proof. By Theorem 10, M is almost chainable. Hence, it follows from Theorem 5 that every nondegenerate proper subcontinuum of M is a pseudo-arc.

Remark. The pseudo-arc **(1**; **13)** is the only known example of a continuum which satisfies the hypothesis of Theorem 11. While the pseudo-circle **(2)** is not known to be homogeneous, it is described with circular chains of open disks and each of its nondegenerate proper subcontinua is a pseudo-arc. It would be interesting to know whether a plane continuum is a pseudo-circle if it is circularly chainable, hereditarily indecomposable, and different from a pseudo-arc. This is suggested by Bing's result that a continuum is a pseudo-arc if it is linearly chainable and hereditarily indecomposable **(2)**.

THEOREM 12. *If the tree-like continuum M is k-branched and nearly homogeneous, then M is indecomposable.*

Proof. Rosen has shown that every k-branched continuum is irreducible about some k points **(15)**, and such an irreducible continuum is indecomposable if it is nearly homogeneous **(7)**.

Remark. Since every tree-like continuum is hereditarily unicoherent **(6)**, it follows from a result by F. B. Jones that every homogeneous tree-like continuum is indecomposable **(11)**. However, it is necessary in Theorem 12 to require that M be k-branched, or at least that it be k-junctioned, as there exists a dendron which is nearly homogeneous **(9)**.

THEOREM 13. *If the indecomposable tree-like continuum M is k-junctioned and nearly homogeneous, then M is almost chainable.*

The following definition and lemma will be used in the proof of Theorem 13.

Definition. A junction link L of a tree T is said to be a *free junction link* of T if there does not exist a linear chain in T which contains L and has two junction links of T different from L as end links.

[6]A forthcoming paper by R. H. Bing will include an affirmative answer to this question. Hence the hypotheses of Theorems 10 and 11 can be weakened accordingly.

LEMMA 13.1. *If the tree-like continuum M is k-junctioned and nearly homogeneous, U is an open subset of M, and ϵ is a positive number, then there exists an ϵ-tree T which covers M and contains only k junction links such that some free junction link of T is a subset of U.*

Proof of Lemma 13.1. It is easy to see that each tree different from a linear chain has a free junction link. For each positive integer i, let T_i be a $1/i$-tree covering M such that T_i has exactly k junction links, and let K_i be a free junction link in T_i. Some subsequence of the sequence $K_1, K_2, K_3, \ldots$ converges to a point p. For convenience, suppose that $K_1, K_2, K_3, \ldots$ converges to p. There is a homeomorphism f of M onto itself that carries p into a point of U. Hence for infinitely many integers i, $f(K_i)$ is a subset of U. From this and the uniform continuity of f, it follows that, for some integer n, $f(K_n)$ is a subset of U and each link of T_n has an image, under f, with a diameter less than ϵ. The collection consisting of all images, under f, of links of T_n is a tree T satisfying the requirements of the conclusion of Lemma 13.1.

Proof of Theorem 13. Suppose that M fails to be almost chainable. There exists a positive number ϵ such that every ϵ-tree covering M has at least k junction links and such that no ϵ-covering of M satisfies the requirements for M to be almost chainable. It follows from the indecomposability of M that there exists a collection W consisting of $2k$ disjoint subcontinua of M such that, for each element X of W, each point of M is within a distance $\epsilon/2$ of X. Let δ be a positive number less than $\epsilon/2$ such that no two continua of W are within a distance δ of each other. Let G be a δ-tree which covers M and has only k junction links. No two continua of W intersect the same link of G, so there exist at least k continua of W that do not intersect a junction link of G. From the supposition that M fails to be almost chainable, it follows that no branch of G covers a continuum of W. Now by induction on k, it can be shown that, for any k linear chains in G each of which has two junction links of G as end links, one such chain must contain at least three junction links of G. Hence there exist two continua H and K of W and a linear chain $C(L_1, L_2, \ldots, L_j, \ldots, L_n)$ in G such that: (1) no link of C is a junction link of G; (2) H is covered by the linear chain $(L_1, L_2, \ldots, L_j)$; and (3) K is covered by the linear chain $(L_{j+1}, L_{j+2}, \ldots, L_n)$. By Lemma 13.1, there exists a tree G', covering M such that: (1) G' is a refinement of G; (2) G' has exactly k junction links; and (3) some free junction link R of G' is a subset of L_j and is not a subset of any other element of C. Let A denote the collection of all elements X of G' such that some linear chain in G' has both R and X as links and has no more than one link that intersects $L_1 + L_n$. There are two cases to consider.

Case 1. One of the sets L_1 and L_n, say L_1, does not intersect an element of A. Let r be the least positive integer such that L_r contains an element of A that is not in L_{r+1}. For each $i(r < i < n)$, let K_i denote the sum of all elements

of A that lie in L_i. Now, since K is covered by the linear $\epsilon/2$-chain $(L_{j+1}, L_{j+2}, \ldots, L_n)$ and every point of M is within a distance $\epsilon/2$ of K, it follows that every point of M is within a distance ϵ of some link of the linear ϵ-chain $(K_r, K_{r+1}, \ldots, K_{n-1})$. However, since no element of $G' - A$ intersects one of the sets $K_r, K_{r+1}, \ldots, K_{n-2}$, this is contrary to the supposition that M fails to be almost chainable.

Case 2. Each of the sets L_1 and L_n intersects an element of A. There exist linear chains C_1 and C_2 in G' such that C_1 is irreducible from R to L_1 and C_2 is irreducible from R to L_n. Let B denote the collection of all links of G' that lie in a branch of G' that starts at R. From the supposition that M fails to be almost chainable, it follows that neither L_1 nor L_n intersects an element of B. For each i $(1 < i < n)$, let L_i' denote the sum of the elements of the collection $B + C_1 + C_2$ that lie in L_i but not in L_{i+1}. Let G'' denote the collection consisting of $L_2', L_3', \ldots, L_{n-1}'$ and the elements of $G' - (B + C_1 + C_2)$. Then G'' is an ϵ-tree covering M, and each junction link of G'' contains a junction link of G'. However, unless L_j' contains a junction link of G' different from R, L_j' is not a junction link of G''. Hence G'' has no more than k-1 junction links. This involves a contradiction as ϵ was chosen so that every ϵ-tree covering M would have at least k junction links.

THEOREM 14. *If the tree-like continuum M is k-branched and nearly homogeneous, then M is almost chainable.*

Proof. A k-branched continuum is at most $(k-2)$-junctioned. Hence Theorem 14 follows from Theorems 12 and 13.

THEOREM 15. *If the tree-like continuum M is k-junctioned and homogeneous, then every nondegenerate proper subcontinuum of M is a pseudo-arc.*

Proof. As observed in the remark following Theorem 12, M is indecomposable. Hence it follows from Theorems 5 and 13 that every nondegenerate proper subcontinuum of M is a pseudo-arc.

COROLLARY. *If the tree-like continuum M is k-branched and homogeneous, then every nondegenerate proper subcontinuum of M is a pseudo-arc.*

Remark. By slight modifications of the arguments, it can be shown that Theorems 5 and 11 and the above corollary hold for a weaker type of homogeneity where, for each point p in the continuum M and each nondegenerate subcontinuum K of M, there is a homeomorphism of M onto itself that carries p into a point of K.

REFERENCES

1. R. H. Bing, *A homogeneous indecomposable plane continuum*, Duke Math. J., *15* (1948), 729–742.
2. ——— *Concerning hereditarily indecomposable continua*, Pacific J. Math., *1* (1951), 43–51.
3. ——— *Snake-like continua*, Duke Math. J., *18* (1951), 653–663.

4. ——— *Each homogeneous nondegenerate chainable continuum is a pseudo-arc*, Proc. Amer. Math. Soc., *10* (1959), 345–346.
5. ——— *A simple closed curve is the only homogeneous bounded plane continuum that contains an arc*, Can. J. Math., *12* (1960), 209–230.
6. R. H. Bing and F. B. Jones, *Another homogeneous plane continuum*, Trans. Amer. Math. Soc., *90* (1959), 171–192.
7. C. E. Burgess, *Some theorems on n-homogeneous continua*, Proc. Amer. Math. Soc., *5* (1954), 136–143.
8. ——— *Homogeneous continua*, Summary of Lectures and Seminars, Summer Institute on Set Theoretic Topology (Madison, 1955), 73–76.
9. ——— *Continua and various types of homogeneity*, Trans. Amer. Math. Soc., *88* (1958), 366–374.
10. ——— *Chainable continua and indecomposability*, Pacific J. Math., *9* (1959), 653–659.
11. F. B. Jones, *Certain homogeneous unicoherent indecomposable continua*, Proc. Amer. Math. Soc., *2* (1951), 855–859.
12. E. E. Moise, *An indecomposable plane continuum which is homeomorphic to each of its nondegenerate subcontinua*, Trans. Amer. Math. Soc., *63* (1948), 581–594.
13. ——— *A note on the pseudo-arc*, Trans. Amer. Math. Soc., *64* (1949), 57–58.
14. R. L. Moore, *Foundations of point set theory*, Amer. Math. Soc. Colloquium Publications, *13*, 1932.
15. Ronald H. Rosen, *On tree-like continua and irreducibility*, Duke Math. J., *26* (1959), 113–122.
16. R. H. Sorgenfrey, *Concerning triodic continua*, Amer. J. Math., *66* (1944), 439–460.

University of Utah

Published Spring 1961

REPRESENTATION THEORY OF THE SYMMETRIC GROUP

BY G. DE B. ROBINSON

This book is devoted to a study of the linear representations, both ordinary and modular, of the symmetric group $\mathfrak{S}_n$, which has come to play an important role in many different contexts. A systematic use of Alfred Young's 'tableau' approach yields constructions which are straightforward and easily understood. An important feature of the book is the evidence which it provides that a modification of the inducing process is much to be desired in the modular case. Mathematical Expositions Series, No. XII. 224 pages 6 × 9 inches $6.00

Soon to be published

GEOMETRY OF THE COMPLEX NUMBERS

BY HANS SCHWERDTFEGER

Mathematical Expositions Series No. XIII

About 216 pages Probably $7.50

Other recent books in the Mathematical Exposition Series

PARTIAL DIFFERENTIAL EQUATIONS

BY G. F. D. DUFF

x + 248 pages 6 × 9 inches $6.50

VARIATIONAL METHODS FOR EIGENVALUE PROBLEMS:

An Introduction to the Methods of Rayleigh, Ritz, Weinstein, and Aronszajn

BY S. H. GOULD

xiv + 179 pages 6 × 9 inches $6.00

DIFFERENTIAL GEOMETRY

BY ERWIN KREYSZIG

xvi + 352 pages 6 × 9 inches $8.50

UNIVERSITY OF TORONTO PRESS

s,
h
s.
ls
d.
o-
e
s,
0

60

60

R

z,

00

50